Distr.
RESTRINGIDA

LC/R.2100
31 de octubre de 2002

ORIGINAL: ESPAÑOL

C E P A L

Comisión Económica para América Latina y el Caribe

9º TALLER REGIONAL

DISEÑO Y CONSTRUCCIÓN DE LOS MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES

Lima, Perú, 17 al 19 de junio de 2002

900028475 - BIBLIOTECA CEPAL

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de la exclusiva responsabilidad de los autores y pueden no coincidir con las de la Organización.

02-11-867

ÍNDICE

Página

PRESENTACIÓN 3

AGENDA 5

PRINCIPALES RESULTADOS Y LOGROS DEL PROGRAMA MECOVI *(BID - Banco Mundial - CEPAL)* 9

RESUMEN Y CONCLUSIONES 21

LOS CENSOS DE POBLACIÓN Y VIVIENDA Y LOS MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES. *Fernando Medina H., CEPAL.* 29

ASPECTOS ESTADÍSTICOS ASOCIADOS A LA CONSTRUCCIÓN DEL MARCO DE MUESTREO. *Fernando Medina H., CEPAL* 39

ASPECTOS ESTADÍSTICOS ASOCIADOS A LA CONSTRUCCIÓN DEL MARCO DE MUESTREO. *Alberto Padilla Trejo, INEI - PERÚ* 63

SISTEMA DE INFORMACIÓN PARA LA SELECCIÓN DE LA MUESTRA. *Juan Valverde Quezada, INEI - PERÚ* 73

DISEÑO DE UN SISTEMA PARA LA SELECCIÓN DE MUESTRAS. *INEI - PERÚ* 83

MARCOS DE ÁREAS PARA ENCUESTAS COMPLEJAS DE VIVIENDAS. *Alfredo Aliaga, CEPAL* 93

INFORME TÉCNICO: EL MARCO DE MUESTREO NACIONAL DE VIVIENDAS Y LA ENCUESTA PERMENENTE DE HOGARES: ESTADO ACTUAL Y SU REFORMULACIÓN. *INDEC - ARGENTINA* 99

INFORME TÉCNICO DEL DISEÑO DE LA MUESTRA MAESTRA DE HOGARES. *Fernando Rivero y Fernando Mollinedo, INE - BOLIVIA* 117

INFORME TÉCNICO SOBRE DISEÑO DE UN MARCO DE MUESTREO DE ÁREAS PARA LAS ENCUESTAS DE HOGARES. *INE - CHILE* 129

DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES. *DANE - COLOMBIA* 143

DISEÑO Y CONSTRUCCIÓN DE MARCOS MUESTRALES PARA ENCUESTAS DE HOGARES. *Giselle Argüello Venegas, INEC – COSTA RICA* 153

MARCOS DE ÁREAS. *Oficina Nacional de Estadística - CUBA* 167

DISEÑOS METODOLÓGICOS DE LA ENCUESTA DE HOGARES DE PROPÓSITOS MÚLTIPLES EHPM. *DGEC – EL SALVADOR* 177

INFORME TÉCNICO DE LAS EXPERIENCIAS EN DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES. *INE - GUATEMALA* 203

DISEÑO MUESTRAL DE LA XXIV ENCUESTA PERMANENTE DE HOGARES. *INE - HONDURAS* 207

INFORME TÉCNICO DE LAS EXPERIENCIAS DE MÉXICO. *INEGI - MÉXICO* 219

ALGUNOS ASPECTOS RELACIONADOS CON LA MUESTRA DE LAS ENCUESTAS DE HOGARES. *Roberto Castillo, DEC - PANAMÁ* 251

DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES. *INEI - PERÚ* 263

MARCO MUESTRAL DE PROPÓSITOS MÚLTIPLES Y MUESTRA MAESTRA. *BANCO CENTRAL - REPÚBLICA DOMINICANA* 291

DISEÑO DE LA MUESTRA DE LA ENCUESTA CONTINUA DE HOGARES. *INE - URUGUAY* 299

MUESTRA MAESTRA CENSO 1990 V/S MUESTRA MAESTRA CENSO 2001. *INE - VENEZUELA* 319

LISTA DE PARTICIPANTES. 325

PRESENTACIÓN

La reducción de los elevados niveles de pobreza e inequidad social que prevalecen en la mayoría de los países de América Latina y el Caribe constituye una de las preocupaciones prioritarias de los gobiernos de la región. La necesidad de evaluar diferentes aspectos de la gestión pública en materia de política social aparece como una actividad cada vez más protagónica en las agendas nacionales, particularmente ante el difícil panorama que en materia de bienestar presentaron los primeros años del presente decenio, donde al débil desempeño mostrado por las economías de la región se sumaron aumentos en las tasas de pobreza en un número importante de países.

Para apoyar el cumplimiento de este objetivo se requiere contar con sistemas de información confiables y oportunos, que permitan conocer, analizar y dar seguimiento a la evolución de los niveles de bienestar de los distintos grupos sociales. En este aspecto, las encuestas de hogares se han consolidado como una herramienta prácticamente indispensable para conocer las repercusiones de la acción pública y la coyuntura económica en el nivel de vida de las personas.

En los años recientes, la mayoría de los países han concretado diversas iniciativas en el ámbito de los sistemas de información que han permitido consolidar los programas de encuestas de hogares. No obstante, a pesar de los logros alcanzados, aún se perciben rezagos que deben ser subsanados, vinculados principalmente a la calidad de la información, así como a las restricciones de cobertura geográfica, al contenido temático de los cuestionarios, a la periodicidad con la que se realizan las encuestas y la calidad de la información que se pone a disposición de los usuarios.

En este contexto, el Banco Interamericano de Desarrollo (BID), el Banco Mundial y la Comisión Económica para América Latina y el Caribe (CEPAL), con la participación directa de los países de la región, han venido desarrollando desde el año 1996 el "Programa para el Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe" (MECOVI). El objetivo central de este Programa es apoyar a los países en la tarea de generar información adecuada y de alta calidad acerca de las condiciones de vida de los habitantes de la región, en cuanto a su contenido, alcance, confiabilidad, actualidad y relevancia para el diseño y evaluación de políticas.

Entre sus actividades, el Programa contempla la realización –bajo la coordinación de la CEPAL– de un conjunto de talleres, pensados como un espacio de reflexión y análisis en el cual los países intercambian sus experiencias y debaten sobre temas de interés común, contribuyendo a fortalecer la capacidad institucional para implementar sistemas de encuestas de hogares y a utilizar eficientemente la información que de ellos se deriva.

Entre 1998 y 2001 se realizaron 8 talleres regionales, en los cuales se abordaron los siguientes temas: Planificación y Desarrollo de Encuestas de Hogares para la Medición de las Condiciones de Vida; Medición del Ingreso en las Encuestas de Hogares; Medición del Gasto en las Encuestas de Hogares; La medición de la Pobreza: El Método de las Líneas de Pobreza; e Indicadores sobre el Desarrollo Social. En todos los casos, a la reflexión teórica y conceptual se incorporaron las experiencias nacionales, que permitieron evaluar las bondades y limitaciones de las distintas opciones metodológicas e identificar áreas de oportunidad para continuar perfeccionando los mecanismos de generación de información. Debido a la importancia que el programa MECOVI le confiere al perfeccionamiento de las fuentes de información, el séptimo Taller Regional se centró en discutir los aspectos conceptuales y técnicos asociados al "Diseño e Implementación de un Sistema Integrado de Encuestas de Hogares (SIEH)". Por su parte, la agenda del octavo taller respondió al interés manifestado por los países beneficiarios y las instituciones copatrocinadoras de evaluar, a partir de las experiencias nacionales, los aportes del programa MECOVI en ámbitos como el fortalecimiento institucional, la mayor disponibilidad de información sobre los determinantes de la pobreza, y la consolidación de los Sistemas Integrados de Encuestas de Hogares como un patrimonio de los Institutos Nacionales de Estadística (INE).

Finalmente, el noveno Taller Regional del MECOVI se realizó en la ciudad de Lima, Perú, en conjunto con el Instituto Nacional de Estadística e Informática (INEI) de ese país. El tema central para este encuentro, el "Diseño y Construcción de Marcos de Muestro para las Encuestas de Hogares", sirvió para fomentar la generación de sinergias institucionales en torno a la realización de los Censos de Población y Vivienda en los países de la región.

Al igual que los materiales elaborados con oportunidad de los talleres anteriores, toda la documentación contenida en este libro se encuentra también a disposición de las personas e instituciones interesadas en el sitio de la CEPAL en Internet[1].

[1] La dirección en Internet en la que se encuentran disponibles los documentos es:
http://www.eclac.cl/deype/noticias/proyectos/7/5497/mecovi/ACTIVIDADREGIONAL/TALLERES/TALLERES.html

PROGRAMA PARA EL MEJORAMIENTO DE LAS ENCUESTAS Y LA MEDICION DE LAS CONDICIONES DE VIDA EN AMERICA LATINA Y EL CARIBE (MECOVI)
BID - BANCO MUNDIAL - CEPAL

9° TALLER REGIONAL
DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO
PARA LAS ENCUESTAS DE HOGARES
Lima, Perú, 17 al 19 de junio de 2002

INEI **CEPAL**

Agenda

Lunes 17	
09:00 – 09:30	Registro de participantes
09:30 – 09:50	Palabras de bienvenida *(INEI, CEPAL)*
09:50 – 10:00	Presentación de la Agenda y de la organización del Taller *(CEPAL)*
TEMA 1	**Los Censos de Población y Vivienda y los Marcos de Muestreo utilizados en las encuestas en hogares**
10:00 – 13:00	1.1 Relación entre la Cartografía Censal y el marco de muestreo utilizado para la selección de la muestra de viviendas 1.2 Utilización del material censal para la construcción de los marcos de muestreo 1.3 Preparación y adaptación de los mapas censales para la elaboración del marco de muestreo
11:00 – 11:15	Café
13:00 – 15:00	Almuerzo
TEMA 2	**Diseño y Construcción del Marco de Muestreo**
15:00 – 18:00	2.1 Importancia de los Marcos de Áreas para la selección de Viviendas 2.2 Requerimientos de información, dominios de estudio e inferencia estadística 2.3 Los Marcos Maestros de Muestreo y los diseños de muestras maestras 2.4 Procedimiento para la formación de las Unidades del Marco de Muestreo a) Identificación de límites b) Numeración e identificación de las unidades de marco c) Asignación de medidas de tamaño 2.5 Tamaño de las Unidades Primarias de Muestreo y su efecto en la varianza de los estimadores 2.6 Tamaño óptimo de las unidades secundarias y terciarias de muestreo (segmentos o conglomerados de viviendas) y su influencia en la varianza de los estimadores
16:30 – 16:45	Café

Martes 18	
TEMA 3	**Aspectos estadísticos asociados a la construcción del marco de muestreo**
09:00 – 18:00	3.1 Procedimientos para la estratificación del Marco de Muestreo a) Ventajas de la estratificación del marco b) Métodos de estratificación: criterios administrativos y procedimientos estadísticos 3.2 Selección de las unidades de muestreo y requerimientos de información a) Unidades de selección autorepresentadas b) Tasas de selección para áreas de listado y viviendas c) Muestras autoponderadas d) Selección de réplicas (submuestras) 3.3 Métodos para la asignación de la muestra 3.4 Procedimientos para la selección de la Muestra 3.5 Diseños en tres etapas: ventajas y desventajas 3.6 Influencia del nuevo marco muestral sobre los resultados de la encuesta 3.7 Muestras de panel de hogares y viviendas 3.8 Uso de proyecciones de población para generar estimaciones
11:00 – 11:15	Café
13:00 – 15:00	Almuerzo
16:30 – 16:45	Café
Miércoles 19	
TEMA 4	**Aspectos operativos para la construcción del marco de muestreo**
09:00 – 13:00	4.1 Creación de Áreas de Listado (diferencias entre áreas urbanas y rurales) 4.2 Procedimientos operativos para la elaboración de listados de viviendas a) Reglas prácticas para efectuar recorridos en zonas altamente pobladas y edificios b) Segmentación al interior de las áreas de listado c) Tratamiento a viviendas vacías, en construcción, puertas falsas y transitorias 4.3 Procedimiento para la actualización de las áreas de listado 4.4 Actualización permanente del marco de muestreo y de la muestra a) Procedimientos para alcanzar un adecuado nivel de eficiencia en el trabajo de campo b) Asegurar que la muestra refleje de modo apropiado los cambios que ocurren en el terreno 4.5 Tratamiento de viviendas adicionales, omitidas y desaparecidas a) Nuevas construcciones: el método del intervalo semiabierto

	b) La conveniencia de sustituir viviendas en muestra c) Control del tamaño de muestra, tratamiento de la no respuesta y ajuste en los factores de expansión 4.6 Tratamiento de estratos especiales de viviendas
11:00 – 11:15	Café
13:00 – 15:00	Almuerzo
TEMA 5	**Sistema de información para la selección de la muestra**
15:00 – 17:30	5.1 Cartografía digitalizada 5.2 Selección automatizada de las unidades de marco 5.3 Control automatizado de los listados de viviendas 5.4 Cálculo y ajuste automatizado de los factores de expansión
16:30 – 16:45	Café
17:30 – 17:50	Síntesis y Conclusiones
17:50 – 18:00	Clausura

PRINCIPALES RESULTADOS Y LOGROS DEL PROGRAMA MECOVI

(diciembre 1996-junio 2002)

BID-BANCO MUNDIAL-CEPAL

ÍNDICE

Página

A. Actividades en Países Participantes 11
B. Actividades Regionales 17
C. Incorporación de Nuevos Países 18
D. Evaluación 18
E. Coordinación 19
F. Conferencia Estadística de las Américas (CEA) 19

1.1 A continuación se presenta un breve resumen de los reportes de avances de actividades por país preparados para el Comité Directivo (Steering Committee) y el Consejo Asesor (Advisory Board) del Programa MECOVI.

1.2 El Programa MECOVI tiene dos componentes principales: (a) actividades en el ámbito de países participantes; y (b) actividades a escala regional (países de América Latina y el Caribe).

A. Actividades en Países Participantes

1.3 A la fecha los países participantes, a través de las oficinas nacionales de estadística, son Argentina, Bolivia, El Salvador, Guatemala, Nicaragua, Paraguay y Perú. Se están finalizando las actividades de coordinación para consolidar la incorporación al Programa de Ecuador, Honduras y República Dominicana. La asistencia técnica otorgada por el Programa a los institutos nacionales de estadística de estos países ha mejorado las encuestas de hogares en términos de su cobertura temática y geográfica. En cuanto al contenido de las encuestas de hogares, se ha expandido la cobertura temática de los cuestionarios tradicionales para incorporar módulos nuevos en algunos o todos los países beneficiarios, incluyendo la medición de la nutrición de los niños (antropometría), remesas familiares, actividad agropecuaria, uso del tiempo de los miembros del hogar, capital social, acceso a programas sociales, seguridad ciudadana, y manejo de los choques adversos de los hogares (ver Cuadro 1). Se ha ampliado la cobertura geográfica de las encuestas al área rural por primera vez en algunos países (Paraguay, Bolivia), en donde justamente se presenta una mayor incidencia y severidad de la pobreza.

1.4 *Asistencia técnica.* El Programa MECOVI ha otorgado asistencia técnica, por medio de consultores nacionales e internacionales expertos en encuestas de hogares, sobre los siguientes aspectos: i) actualización de la cartografía y el diseño de la muestra; iii) desarrollo de cartografía digitalizada; iv) organización del trabajo de campo; v) organización de la entrada de datos de modo descentralizado; vi) desarrollo de los principales criterios de consistencia y control de la calidad en el campo; vii) organización de la base de datos; viii) procesamiento y análisis de la información.

1.5 *Interacción entre usuarios y productores de información.* Se han promovido mecanismos para lograr un mayor estrechamiento de la relación entre productores y usuarios de encuestas de hogares a fin de asegurar que las encuestas capten la información requerida por el país. Para tal efecto, en todos los países participantes se ha conformado un Comité Interinstitucional, compuesto por representantes de diversos organismos públicos y privados, productores y usuarios de encuestas de hogares.

1.6 *Capacitación.* Las oficinas de estadística de los países participantes han organizado talleres de trabajo y cursos de capacitación, con la participación de productores y usuarios de encuestas, para discutir tanto aspectos de metodología y como de resultados de los estudios realizados con la información de las encuestas de hogares. Los temas de los cursos han sido muestreo, uso de software para procesar encuestas, aspectos conceptuales sobre condiciones de vida, etc.

1.7 *Actividades por país.* A continuación se presentan los cuadros resumen por país ordenados cronológicamente en base al inicio de las actividades del Programa.

1.8 *Paraguay.*

Organismo Ejecutor	Dirección General de Estadísticas, Encuestas y Censos (DGEEC)
Inicio de actividades	Las actividades de MECOVI se iniciaron a fines de 1996.
Operación de campo	Encuesta Integrada de Hogares (EIH) a nivel nacional (urbana y rural) de agosto de 1997 a agosto de 1998. Encuesta Permanente de Hogares (EPH), nacional, de agosto a diciembre 1999. Actualmente la EIH 2000-01 está siendo procesada.
Cuestionarios	Cuestionarios de hogares, a centros poblados rurales, de precios, y sobre antropometría.
Base de datos	La base de datos de la EIH 1997-98, y EPH 1999 está disponible para los usuarios.
Asistencia técnica	La asistencia técnica ha contribuido a mejorar: i) el contenido temático de la encuesta; ii) la actualización de la cartografía y el diseño de la muestra; iii) el desarrollo de la cartografía digitalizada; iv) la organización del trabajo de campo; v) la entrada de datos de modo descentralizado; vi) el desarrollo de los principales criterios de consistencia y control de la calidad en el campo; vii) la organización de la base de datos; vii) el procesamiento y análisis de la información.
Capacitación y talleres	Se han organizado cursos de entrenamiento en el manejo de software para el procesamiento de datos, muestreo y análsis de encuestas. Además, los técnicos se han visto beneficiados mediante la capacitación en el terreno con la presencia de los consultores nacionales e internacionales, y los talleres de trabajo realizados.
Estudios y publicaciones	Se han realizado publicaciones y estudios utilizando la información de las encuestas de 1995, 1996 y 1997/98: i) edición de documentos de Divulgación Popular (No 1 y 2); ii) elaboración del estudio sobre el perfil de condiciones de vida en 1995; iii) perfil de condiciones de vida en 1996; iv) estudios descriptivos sobre la situación de la fuerza laboral en 1996, perfil de la salud 1996, y perfil educativo 1996; v) perfil de condiciones de vida 1997/98; vi) revista "Economía & Sociedad" números 1 al 5 y, vii) folletos informativos.
Comité Interinstitucional	El Comité está formado por representantes de: Secretaría Técnica de Planificación de la Presidencia de la República, Banco Central de Paraguay, Centro de Documentación y Estudio (CDE), Universidad Nacional de Asunción, y la DGEEC.
Fondo de Estudios	El Fondo de Estudios se puso en marcha en 1999, utilizando la información de la EIH 1997-98. Se han financiado tres estudios en 1999. La segunda ronda del Fondo de Estudios ha concluido. Los estudios preparados han sido publicados como parte de la revista "Economía & Sociedad".
Evaluación	Durante 1998 se realizó una evaluación de los cuestionarios ejecutados (hogares y comunitario). En el IV trimestre de 2000 se evaluó, con excelentes resultados, el proceso de implementación de la asistencia técnica otorgada.

1.9 *Perú.*

Organismo Ejecutor	Instituto Nacional de Estadística e Informática (INEI)
Inicio de actividades	En Perú se iniciaron las operaciones del Programa MECOVI en Febrero de 1997.
Operación de campo	Se han completado las rondas del IV trimestre de 1997, del II, III y IV trimestres de 1998, del I, II, III y IV trimestres de 1999, y del I, II, III y IV trimestres de 2000 de la "Encuesta Nacional de Hogares" (ENAHO) con cobertura nacional (urbana y rural). Las encuestas correspondientes a 2001 están siendo procesadas.
Cuestionarios	Cuestionarios de hogares, a comunidades rurales, y de precios rurales.
Base de datos	La base de datos de las ENAHO del IV trim. 1997, los trimestres I, II, III y IV de 1998, 1999 y 2000 están disponibles para los usuarios.
Asistencia técnica	La asistencia técnica ha estado orientada a: I) mejorar el contenido temático del cuestionario mediante el análisis de las principales variables usando la ENAHO 1996 (gasto, ingreso, empleo, acceso a servicios sociales básicos); ii) el diseño de la muestra; iii) mejorar los métodos de capacitación de los encuestadores y supervisores (de modo descentralizado); iv) el desarrollo de sistemas de control de calidad en todas las etapas de la encuesta, incluyendo un banco de preguntas y respuestas para compartir los problemas y las soluciones sobre aspectos de la encuesta en el Intranet del INEI; v) mejoramiento del análisis y procesamiento de los datos.
Capacitación y talleres	Se ha reforzado las capacidades de los técnicos del INEI y los usuarios de las encuestas mediante cursos nacionales de capacitación en las siguientes áreas: manejo de software, uso de la base de datos y aplicaciones de la ENAHO, muestreo, análisis de datos. Estos se han complementado mediante el desarrollo de talleres y seminarios metodológicos, a través de la capacitación en el terreno con la presencia de los consultores nacionales e internacionales y mediante cursos y talleres regionales. El INEI ha organizado dos de dichos talleres: el primero, para discutir metodologías utilizadas en común en los países participantes en el MECOVI; el segundo, para examinar el uso de información de encuestas de hogares de condiciones de vida en conjunto con información de censos.
Estudios y publicaciones	Se han elaborado estudios sobre el perfil de la pobreza y distribución del ingreso en el Perú con base a la información de la ENAHO 1996, 1997 y 1998. Asimismo, se han realizado un estudio para definir las líneas de pobreza en Perú y estudios sobre el perfil de condiciones de vida en 1997 y 1998. Además se han hecho publicaciones descriptivas de los sectores sociales (educación, vivienda, salud, empleo, etc.).
Comité Interinstitucional	El Comité está formado por representantes de: Ministerios de la Presidencia, de Trabajo, de Salud, de Educación, de Transportes, Vivienda y Construcción, Promoción de la Mujer y Desarrollo Humano, Economía; Banco Central de Reserva; Fondo de Compensación para el Desarrollo Social (FONCODES); Instituto Nacional de Salud; Universidad del Pacífico; Grupo de Análisis para el Desarrollo (GRADE); Asociación Laboral para el Desarrollo (ADEC-ATC) e INEI.
Fondo de Estudios	El uso del Fondo de Estudios fue activado en 1999, financiando 16 proyectos de investigación. En la segunda convocatoria se han seleccionado 16 nuevos estudios.
Evaluación	Durante 1998 se realizó una evaluación de los cuestionarios ejecutados (hogares y comunitario), en el IV trimestre de 2000 se evaluó, con excelentes resultados, el proceso de implementación de la asistencia técnica otorgada.

1.10 *El Salvador.*

Organismo Ejecutor	Dirección General de Estadística y Censos (DIGESTYC).
Inicio de actividades	Septiembre de 1997.
Operación de campo	La operación de campo de la Encuesta de Hogares de Propósitos Múltiples (EHPM), a nivel nacional (urbano y rural), se ejecutaron entre mayo - diciembre de 1998 y febrero - diciembre 1999. La encuesta de 2000 está siendo procesada, en tanto que la ejecución de la de 2001 se vio retrasada como consecuencia de los terremotos de enero y febrero de 2001.
Cuestionarios	Cuestionarios de hogares, comunitario, precios rurales, y antropometría.
Base de datos	Las bases de datos de la EHPM 1998 y 1999 están disponibles.
Asistencia técnica	La asistencia técnica ha estado dirigida a: i) actualización del marco muestral durante el periodo septiembre 1997- marzo 1998; ii) mejorar el contenido temático de los cuestionarios; iii) diseño de la muestra; iv) organizar el trabajo de campo y la entrada de datos; v) sistematizar métodos de control de calidad; vi) definir un plan de trabajo orientado a la digitalización de la cartografía; vii) uso de software para el manejo de bases de datos (SPSS, ISSA); y, viii) recolección y manejo de información antropométrica.
Capacitación y talleres	Se han realizado talleres de trabajo de tipo metodológico para discutir la definición de los cuestionarios, muestra y control de calidad, con la participación de consultores nacionales e internacionales. Se han realizado cursos en el uso de software para manejo de datos, sobre antropometría, y análisis de encuestas de hogares.
Estudios y publicaciones	Estimados mensuales de la tasa de desempleo. Estudios descriptivos sobre nutrición, condiciones de vida y tabulados generales.
Comité Interinstitucional	El Comité está conformado por representantes de: Ministerios de Salud, de Educación, de Relaciones Exteriores, y de Trabajo; Banco Central de Reserva; FUSADES; Fondo de Inversión Social (FISDL); Universidad José Simeón Cañas; DIGESTYC.
Fondo de Estudios	El uso del Fondo de Estudios fue activado en 1999 con la selección de cuatro estudios. Actualmente se está convocando a una segunda selección de estudios.
Evaluación	Durante 1998 se realizó una evaluación de los cuestionarios ejecutados (hogares y comunitario). En el IV trimestre de 2000 se evaluó, con excelentes resultados, el proceso de implementación del programa.

1.11 *Argentina.*

Organismo Ejecutor	Instituto Nacional de Estadística y Censos (INDEC).
Inicio de actividades	Noviembre 1998.
Operación de campo	Ejecución de la "Encuesta Permanente de Hogares" (EPH), dos veces por año. Una versión revisada de esta encuesta será ejecutada de modo continuo. Se está procesando el Encuesta de Condiciones de Vida 2001.
Cuestionarios	Cuestionario de hogares básico (ingreso y empleo). Cuestionario ampliado sobre condiciones de vida (a ser implementado en 2001).
Base de datos	Disponibles tres meses después de terminar el trabajo de campo.

Asistencia técnica	En proceso: (i) actualización del marco muestral; (ii) rediseño del cuestionario de la EPH; (iii) estrategia para un sistema integrado de encuestas de hogares; (iv) módulo de seguimiento del consumo; (v) revisión de la medición del ingreso; (vi) indicadores sociodemográficos; (vii)diseño de publicaciones; (viii) indicadores laborales; (ix) seguimiento del consumo; y (x) armonización de los conceptos y definiciones de la encuesta y los registros administrativos.
Capacitación y talleres	Se han realizado dos cursos nacionales sobre muestreo, un curso sobre indicadores sociales, y dos talleres nacionales. En marzo del 2000 se llevó a cabo el tercer curso regional de capacitación en el INDEC con la participación de 35 estudiantes de 18 países. En abril de 2001 se realizó el quinto curso regional del MECOVI en el INDEC, se contó con la asistencia de 30 técnicos de 18 países. En abril de 2002 se llevó a cabo el séptimo curso regional del MECOVI en el INDEC, en el cual participaron 30 técnicos de 18 países.
Estudios y publicaciones	(i) Estudio de los gastos e ingresos de los hogares (ENGHO 1996-97); (ii) análisis del mercado laboral; y, (iii) educación.
Comité Interinstitucional	En proceso de formación.
Fondo de Estudios	El fondo fue activado en 1999. Actualmente se están financiando 11 propuestas de investigación (6 de investigadores expertos y 5 de estudiantes de posgrado). En la convocatoria de 2000 se financiaron ocho proyectos (cuatro de investigadores expertos y cuatro de estudiantes de posgrado).
Evaluación	En mayo de 2002 se realizó una evaluación de medio término del programa.

1.12 *Nicaragua.*

Organismo Ejecutor	Instituto Nacional de Estadística y Censos (INEC).
Inicio de actividades	Mayo 1998
Operación de campo	"Encuesta de Medición de Niveles de Vida" (EMNV), en áreas urbanas y rurales: mayo – agosto 1998. Encuesta de seguimiento del impacto del Huracán Mitch: 1999. Se finalizó el trabajo de campo de la EMNV-2001, se está procesando la base de datos.
Cuestionarios	Cuestionarios de hogares, comunitario, precios rurales, y antropometría.
Base de datos	La base de datos de la EMNV 1998 y 1999 está disponible para los usuarios
Asistencia técnica	(i) Diseño de cuestionario; (ii) diseño de muestra; (iii) organización del trabajo de campo; (iv) organización de la base de datos; (v) procesamiento y análisis de la información; (vi) generación del agregado de consumo; y (vii) construcción de un mapa de pobreza.
Capacitación y talleres	Entrenamiento en el uso de SPSS. En 1999 se realizaron talleres sobre muestreo y metodología de la encuesta de hogares.
Estudios y publicaciones	Folleto de divulgación popular y las principales tabulaciones. Informe general de resultados de la EMNV 1998-99. Estudio sobre pobreza rural, y un mapa de pobreza..
Comité Interinstitucional	Miembros: Instituto Nacional de Estadística y Censos (INEC); Ministerio de Economía y Desarrollo (MEDE); Ministerio de Educación (MED); Ministerio de Salud (MINSA); Ministerio de Trabajo (MITRAB); Instituto de fomento Municipal (INIFOM); Ministerio de Agricultura y Ganadería (MAG) ; Banco Central de Nicaragua (BCN); Fondo de Inversión Social de Emergencia (FISE); Ministerio de Acción Social (MAS); Programa de las Naciones Unidas para el Desarrollo (PNUD); Fondo de Población de las Naciones Unidas (FNUAP); y, Banco Mundial.
Fondo de Estudios	El fondo será activado en el año 2000.
Evaluación	Se realizó entre octubre y noviembre de 2000, con buenos resultados.

1.13 *Bolivia.*

Organismo Ejecutor	Instituto Nacional de Estadística (INE).
Inicio de actividades	Mayo 1999
Operación de campo	Encuesta Continua de Hogares (ECH), noviembre 1999, y Encuesta Continua de Hogares IV trimestre 2000 y IV trimestre 2001.
Cuestionarios	Cuestionarios de hogares.
Base de datos	Disponible ECH, 1999.
Asistencia técnica	i) Diseño de cuestionario; ii) diseño de muestra; y, iii) organización de trabajo de campo, y (iv) metodología para la consistencia de datos.
Capacitación y talleres	Se han programado cursos y talleres para técnicos del INDEC y usuarios externos.
Estudios y publicaciones	Informe de resultados de la ECH 1999.
Comité Interinstitucional	A ser conformado.
Fondo de Estudios	El fondo de estudios se activará en 2001.
Evaluación	En marzo de 2002 se llevó a cabo una evaluación de medio término del programa.

1.14 *Guatemala.*

Organismo Ejecutor	Instituto Nacional de Estadística (INE).
Inicio de actividades	Octubre de 1999
Operación de campo	"Encuesta Nacional de Condiciones de Vida" (ENCOVI) (julio – noviembre de 2000) terminada. Se está trabajando en el diseño de la Encuesta Nacional de Empleo e Ingresos (ENEI).
Cuestionarios	Cuestionarios de hogares, comunitario, precios, y antropometría.
Base de datos	En proceso de limpieza y consistencia, se tendrá disponible en junio 2001.
Asistencia técnica	Asistencia técnica en el diseño del cuestionario, actualización del marco muestral, y organización de trabajo de campo. Además, en el análisis de la información y procesamiento y manejo de datos de la "Encuesta Nacional de Ingresos y Gastos Familiares" (ENIGFAM, 1998).
Capacitación y talleres	Por ejecutar.
Estudios y publicaciones	Estudio en proceso sobre condiciones de vida utilizando datos de ingreso y gasto de la "Encuesta Nacional de Ingresos y Gastos Familiares" (ENIGFAM) 1997-98.
Comité Interinstitucional	Ministerio de Salud, Ministerio de Trabajo, Banco de Guatemala, FONAPAZ, INCAP/OPS, SEGEPLAN, GTZ, Fundación Soros, CIEN, Universidad del Valle, Universidad Rafael Landívar, UNFPA, PNUD, UNICEF, USAID, FLACSO, DIACO, Instituto Nacional de Estadística.
Fondo de Estudios	A ser activado en 2000.
Evaluación	Se realizará a los dos años de iniciada la ejecución.

B. Actividades Regionales

1.15 Las actividades regionales del Programa MECOVI incluyen: talleres metodológicos, cursos de capacitación y un banco de datos. A la fecha se han realizado las siguientes actividades:

1.16 *Talleres regionales.* El Programa MECOVI con la meta de asegurar la continuidad de las actividades de producción de encuestas de hogares de calidad por parte de las oficinas de estadística, ha promovido actividades para crear una masa crítica de técnicos en encuestas de hogares, productores y usuarios, mediante la organización de talleres de trabajo y cursos de capacitación tanto a escala nacional como regional, y a través de la asistencia técnica en el terreno por medio de los expertos en encuestas. A la fecha se han organizado ocho talleres de trabajo a nivel regional para discutir aspectos metodológicos de las encuestas de hogares. El primero, "Planificación y desarrollo de encuestas de hogares para la medición de condiciones de vida" (1-3 de abril de 1998 en México). El segundo, "Medición de los ingresos del hogar" (10-13 de noviembre de 1998 en Argentina). El tercero, "Medición del gasto de los hogares" (24-28 de mayo de 1999 en México). El cuarto, "La medición de la pobreza: El método de las líneas de pobreza" (16-19 de noviembre de 1999 en Argentina). El quinto taller, "La medición de la pobreza: Métodos y Aplicaciones (continuación)" (6-8 de junio, 2000, en México). El sexto taller, "Indicadores sobre el desarrollo social" (15-16 de noviembre, 2000, en Argentina). El séptimo taller, "Diseño e Implementación de un Sistema Integrado de Encuestas de Hogares" (30 de mayo al 1ero de junio, 2001, en México). El octavo taller se llevó a cabo en Buenos Aires del 28 al 30 de noviembre, 2001, el tema fue "Pasado, Presente y Futuro del Programa MECOVI". En todos los talleres participaron representantes, productores y usuarios de la mayoría de países de la región.

1.17 *Cursos de capacitación.* Se han realizado siete cursos de capacitación regional sobre "Diseño, implementación y análisis de encuestas de hogares", de tres semanas de duración cada uno. El primero del 1-17 de marzo de 1999, y el segundo del 30 de agosto al 17 de septiembre de 1999, en conjunto con el Instituto Nacional de Estadística, Geografía e Informática (INEGI) de México. El tercer curso se realizó en cooperación con el Instituto Nacional de Estadística y Censos (INDEC) de Argentina, en Buenos Aires, del 29 de marzo al 18 de abril de 2000. El cuarto tuvo lugar en el INEGI (4 al 22 de septiembre de 2000). El quinto curso se llevó a cabo en cooperación con el INDEC, del 16 de abril al 4 de mayo de 2001 en Argentina. El sexto curso tuvo lugar en el INEGI (3 al 21 de septiembre, 2001). El séptimo curso se organizó en coordinación con el INDEC, del 8 al 26 de abril, 2002. En los siete cursos se contó con la participación de un total de 230 productores y usuarios de encuestas de la mayoría de los países de la región.

1.18 *Fortalecimiento de la Capacidad de Análisis de la Información.* El Programa MECOVI con la finalidad de potenciar un amplio uso y análisis de la información de las encuestas de hogares generadas en los países, ha implementado un Fondo de Estudios con el objetivo de promover y apoyar la realización de investigaciones por profesionales nacionales. En este contexto, a la fecha en Paraguay se han concluido tres estudios correspondientes a la primera ronda, y nueve de la segunda ronda, así como cuatro estudios en El Salvador, 16 estudios en Perú y 19 en Argentina. Los otros países están en proceso de lanzar la convocatoria.

1.19 *Mejoramiento del Acceso y Difusión de la Información.* Con ayuda de la asistencia técnica los países han mejorado la organización y documentación de las encuestas de hogares. Además un elemento esencial que se ha promovido es el amplio acceso a la información obtenida, tanto a las bases de datos originales como a los resultados estimados. Las oficinas de estadística están proveyendo esta información a los usuarios del gobierno, universidades, investigadores independientes, y otras instituciones sin fines de lucro.

1.20 *Publicaciones.* Se ha brindado apoyo para mejorar el formato y contenido de las publicaciones de resultados de las encuestas, así como para perfeccionar los modos de difusión de las mismas. En esta línea, los países han publicado una serie de documentos tanto de resultados como metodológicos y han sido difundidos en forma impresa (libros, boletines de prensa, reportes y folletos) y electrónica (CD-Rom e Internet). Al respecto, los países participantes se encuentran desarrollando sus respectivas páginas electrónicas para divulgar sus documentos, varios de los cuales están actualmente disponibles en el sitio de la Unidad de Pobreza y Desigualdad del BID (http://www.iadb.org/sds/pov), en donde se encuentran enlaces a las oficinas de estadística que cuentan con páginas electrónicas, así como a la CEPAL y el Banco Mundial.

1.21 *Organización de un Banco de Datos.* Se ha conformado un Banco de Datos que contiene 180 bases de datos de encuestas de hogares de 21 países de la región, disponible de modo inmediato a los usuarios internos del Banco, y con proyección de ampliar su acceso a usuarios externos. Este Banco de Datos incluye las bases de datos de 67 encuestas de hogares que han sido estandarizadas en el formato y documentadas por la CEPAL en función a los antecedentes metodológicos disponibles y al contenido de los archivos.

C. Incorporación de Nuevos Países

1.22 De acuerdo al Plan de Operaciones del Programa MECOVI y siguiendo la recomendación del Comité Directivo de incorporar dos (2) nuevos países por año al Programa, a la fecha, se han incorporado Nicaragua (junio 1998), Bolivia (mayo 1999) y Guatemala (septiembre 1999). Se esta finalizando el proceso de incorporación de Ecuador, Honduras y República Dominicana. Otros países que han expresado su interés en participar en el Programa MECOVI son Costa Rica, Haití, Panamá, y Venezuela. Al respecto, el BID ha aprobado una nueva cooperación técnica a fin de explorar e identificar las necesidades de asistencia técnica, material y financiera para desarrollar un sistema de encuestas de hogares en los países candidatos a ser incorporados en el programa. Asimismo, dicha cooperación tendrá como objetivo identificar fuentes potenciales de financiamiento y desarrollar un plan de acción para hacer efectiva la incorporación de nuevos países al programa en el periodo 2000-2002.

1.23 Ecuador. En el caso de Ecuador el apoyo a las encuestas de hogares se está dando en el marco de un préstamo solicitado por el Gobierno al BID para financiar las actividades del Censo del 2001 y para fortalecer al Instituto Nacional de Estadística y Censos (INEC).

1.24 Honduras. Se han realizado tres misiones exploratorias, la primera del 18 al 21 de septiembre, 2000, la segunda del 14 al 18 de mayo de 2001, y la última del 13 al 16 de agosto de 2001. En conjunto con el Instituto Nacional de Estadística, se elaboró un plan de operaciones. Se espera que la mayoría de los recursos provengan de una operación de préstamo del BID.

1.25 República Dominicana. Se llevó a cabo una misión indagatoria del 10 al 13 de octubre, 2000, se ha preparado una primera propuesta de plan de operaciones. El Banco Mundial ha aprobado recursos iniciales para apoyar el mejoramiento de las encuestas de hogares en el país.

D. Evaluación

1.26 Se llevó a cabo una evaluación de las actividades y resultados del Programa MECOVI en Argentina, Bolivia, El Salvador, Paraguay y Perú. La finalidad de este ejercicio es obtener recomendaciones que permitan mejorar la implementación del Programa en dichos países, y obtener lecciones aplicables en aquellos que se han incorporado al MECOVI más recientemente.

E. Coordinación

1.27 El Comité Directivo es la instancia encargada de dictar el rumbo del Programa, está compuesto por el Vicepresidente del BID, el Vicepresidente para América Latina y el Caribe del Banco Mundial, y el Secretario Ejecutivo de la CEPAL. A la fecha ha tenido tres reuniones, el 20 de febrero de 1997, el 27 de julio de 1998, y el 17 de marzo de 2001.

1.28 Durante la última reunión del Comité Directivo, 17 de marzo de 2001 en la sede de la CEPAL en Santiago, los miembros del Comité coincidieron en subrayar la importancia del trabajo que el Programa ha realizado en sus primeros cuatro años de labores. Además, se comprometieron a seguir apoyando los trabajos de coordinación y monitoreo en las tres instituciones co-patrocinadoras, y a continuar la búsqueda de recursos para apoyar a los nuevos países que se incorporen al MECOVI y para seguir con las actividades regionales del Programa.

1.29 El Consejo Asesor está conformado por expertos en el tema de encuesta de hogares y tiene como función dar consejos técnicos para la ejecución del Programa. A la fecha se ha reunido en dos ocasiones, 17 de enero de 1997 y el 6 de noviembre de 1998.

1.30 El Comité de Coordinación está formado por los coordinadores del Programa MECOVI en el BID, el Banco Mundial y la CEPAL, tiene como función la ejecución de las actividades básicas del Programa.

F. Conferencia Estadística de las Américas (CEA)

1.31 El Programa MECOVI formó parte de la agenda de la primera reunión de la Conferencia Estadística de las Américas (9-11 de mayo de 2001, en la sede de CEPAL en Santiago), y sus actividades quedaron incluidas en el programa de trabajo estadístico internacional para América y el Caribe 2001-2002.

Cuadro 1. Programa MECOVI: Encuestas de Hogares.

Módulos	Argentina (a)	El Salvador (b)			Perú (c) (98-01)				Paraguay (d)			Nicaragua (e)		Bolivia (f)	Guatemala (g)
		98 (h)	99 (m)	00	I	II	III	IV	97/98 (h)	99 (h)	00/01 (i)	98 y 01 (h)	99 (k)	99 IV (h)	00
1. Características del hogar y sus miembros	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
2. Migración	■	■	□	□	□	■	□	□	■	■	■	□	□	■	■
3. Vivienda	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
4. Salud (general)	□	■	■	■	□	■	□	■	■	■	■	■	■	■	■
5. Salud de la mujer	□	■	□	□	□	■	□	□	■	□	■	■	■	■	■
6. Antropometría	□	■	□	□	□	□	□	□	■	□	■	■	□	□	■
7. Educación	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
8. Participación de padres en educación	□	□	■	■	□	□	□	□	□	□	□	□	□	□	□
9. Acceso a programas sociales	□	□	□	□	■	■	□	■	□	□	□	□	□	□	■
10. Empleo e ingreso	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
11. Actividad independiente del hogar	□	□	□	□	□	□	□	□	■	■	■	■	■	□	■
12. Actividad agropecuaria	□	■	■	■	■	■	■	■	■	■	■	■	■	■	■
13. Gastos del hogar	□	■	■	■	□	□	□	■	■	□	□	■	■	■	■
14. Ahorro financiero, compras a crédito	□	□	□	□	□	□	□	□	□	□	□	■	■	□	■
15. Uso del tiempo	□	□	□	□	□	□	□	□	□	□	■	■	□	□	■
16. Remesas	□	■	■	■	□	□	□	□	□	□	□	□	□	□	■
17. Seguridad ciudadana	□	□	■	■	■	□	□	□	□	□	■	□	□	□	■
18. Choques adversos	□	□	□	□	□	□	□	□	□	□	■	□	■	□	■
19. Etnia	□	□	□	□	□	■	□	□	■	■	■	□	□	■	■
20. Participación ciudadana	□	□	□	□	□	□	□	□	□	□	■	□	□	□	■

NOTAS:

(a) Encuesta Permanente de Hogares (EPH). Esta encuesta se encuentra todavía en la fase de reformulación dentro del contexto del Programa MECOVI. Los módulos señalados son los que se implementan como parte de la EPH tradicional en mayo y octubre de cada año.
(b) Encuestas de Hogares de Propósitos Múltiples (EHPM).
(c) Encuesta Nacional de Hogares (ENAHO).
(d) Encuesta Integrada de Hogares (EIH) 1997/98 y Encuesta Permanente de Hogares (EPH) 1999.
(e) Encuesta Nacional de Hogares sobre Medición de Niveles de Vida (EMNV).
(f) Encuesta Continua de Hogares (ECH).
(g) Encuesta Nacional de Condiciones de Vida (ENCOVI).
(h) Bases de datos disponibles para usuarios.
(i) Encuestas en ejecución.
(k) Esta encuesta es una versión modificada de la EMNV 1998, la sección de choques consiste en una serie de preguntas diseminadas a lo largo de los módulos del cuestionario enfocadas al impacto del Huracán Mitch. Esta encuesta sólo se ejecutó en las áreas afectadas por el Huracán Mitch.
(m) En proceso de revisión y consistencia de las bases de datos de las encuestas.

9° TALLER REGIONAL DEL MECOVI

DISEÑO Y CONSTRUCCIÓN DE LOS MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES

RESUMEN Y CONCLUSIONES

ÍNDICE

Página

I.- Introducción 23
II.- Resultados de los Debates 23
Tema 1. Los Censos de Población y los marcos de muestreo en las encuestas de hogares... 23
Tema 2. Diseño y construcción de los marcos de muestreo 24
Tema 3. Aspectos estadísticos asociados a la construcción del marco de muestreo 26
Tema 4. Aspectos operativos para la construcción del marco de muestra 27
Tema 5. Sistema de Información para la selección de la muestra 28

I. Introducción

El **Noveno Taller Regional del Programa MECOVI** se realizó en la ciudad de Lima, Perú del 17 al 19 de junio de 2002. El tema central que orientó los debates fue el **"Diseño y Construcción de Marcos de Muestreo para las Encuestas de Hogares"**, considerando que la realización del programa de Censos de Población y Vivienda del 2000 representaba una buena ocasión para debatir sobre la renovación de los marcos de muestreo de viviendas.

Esta actividad fue organizada por la División de Estadística y Proyecciones Económicas de la CEPAL y el Instituto Nacional de Estadística e Informática (INEI) de Perú, en el marco del Programa para el Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe (MECOVI), copatrocinado por el BID, el Banco Mundial y la CEPAL.

A lo largo de la existencia de los talleres del MECOVI, se ha buscado constantemente que estos sirvan como espacios de reflexión sobre los temas de actualidad que interesan a los países en relación con la realización de encuestas para la evaluación y seguimiento de los niveles de vida. En esta oportunidad se buscó además propiciar un intercambio de ideas y conocimientos entre los técnicos del área de encuestas de hogares –encargados del diseño de encuestas y de la elaboración del marco de muestreo– y los relacionados con los censos –más específicamente, con los responsables de elaborar la cartografía.

El **objetivo general** del taller fue analizar los aspectos técnicos y operativos que intervienen en el diseño y construcción de los marcos de muestreo, concebidos como un insumo fundamental para planear y ejecutar de manera organizada e integrada las encuestas de hogares, conforme a las prácticas de las Oficinas de Estadística de la región.

A su vez, los **objetivos específicos** del taller fueron los siguientes:

a) Analizar la relación entre la elaboración de la cartografía censal y las necesidades de un marco de muestro de viviendas.

b) Evaluar la planeación estratégica del marco de muestreo, en función a las necesidades y prioridades informativas definidas por los países.

c) Profundizar en el análisis técnico de las decisiones a tomar durante el diseño y la construcción del marco de muestreo.

d) Identificar las buenas prácticas que se aplican en la región para la construcción y administración de los marcos de muestreo.

En este taller participaron 32 técnicos de 16 países de América Latina y el Caribe, además de 6 expertos de las agencias internacionales copatrocinadoras del MECOVI. La agenda de trabajo se desarrolló de acuerdo a lo establecido, y la amplitud y profundidad, tanto de las exposiciones como de los debates, permiten concluir que se cumplieron plenamente los objetivos de la reunión.

II. Resultados de los Debates

TEMA 1. Los Censos de Población y Vivienda y los marcos de muestreo utilizados en las encuestas de hogares

En la mayoría de los países de la región, el período actual coincide con la planificación y ejecución de los Censos de Población y Vivienda, que se efectúan generalmente cada diez años. Dado que un buen número de países de América Latina y el Caribe ya ha concluido el ejercicio censal, en tanto que el resto se encuentra en su etapa de preparación, el taller representó la oportunidad para motivar el intercambio de experiencias sobre los aspectos teóricos y operativos vinculados a la elaboración de los marcos de muestreo. En particular, los diálogos entre los técnicos encargados de la elaboración de la cartografía censal y aquellos relacionados con el diseño y ejecución de las encuestas de hogares

contribuyeron, de alguna manera, a mejorar la comunicación y mutua comprensión entre los funcionarios de los Institutos Nacionales de Estadística (INE).

La ejecución de los Censos de Población y Vivienda representa una buena oportunidad para actualizar y renovar el material cartográfico que sirve de base para la construcción del marco de viviendas utilizado en la planeación de las encuestas de hogares. De hecho, para la mayoría de los países, esta es muchas veces la única ocasión en la que se dispone de recursos para ello y es, por tanto, muy importante que la cartografía atienda tanto a las necesidades del censo como a las de las encuestas de hogares. No obstante, en la práctica la coordinación suele ser inadecuada o inexistente, y actividades tales como el diseño de los cuestionarios y la elaboración de la cartografía censal se realizan frecuentemente sin involucrar a los especialistas en encuestas. Como resultado, los productos generados no responden adecuadamente a las exigencias de información que plantea el diseño de muestras para las encuestas de hogares, lo cual implica realizar gastos adicionales posteriormente.

Una adecuada coordinación va más allá de compartir los mapas producidos, y requiere considerar las necesidades del marco muestral en distintos ámbitos. Por ejemplo, la actualización de las edificaciones y viviendas particulares que existen en el país constituye un insumo indispensable para la construcción del marco de muestreo. Por otra parte, que los segmentos censales se acompañen con su delimitación geográfica y con el número total de edificaciones establecidas al interior de cada uno de ellos facilita el diseño físico de las unidades de marco, así como la aplicación de ponderadores y el control de las probabilidades de selección de las viviendas.

Aún cuando en la mayoría de los países la elaboración de la cartografía censal no atiende a las necesidades planteadas por las encuestas de hogares, los mapas censales representan el insumo fundamental para la construcción del marco. Para lograr el total aprovechamiento de la cartografía censal en la selección de viviendas, existen ciertas diferencias de base a ser superadas. Los desafíos más grandes señalados por los participantes se relacionan con el tamaño y la delimitación de los segmentos censales y la rápida desactualización de la cartografía. Respecto a este último punto, se especificó que mientras más tiempo transcurre entre la elaboración de la cartografía censal y la planificación del marco de muestreo, mayores son las inconsistencias que pueden generarse, haciendo más complejas las labores de diseño y construcción del marco de muestreo y generando costos adicionales que los INE no siempre pueden asumir.

Por otra parte, la construcción de marcos maestros de muestreo en los países se planteó como una alternativa pertinente para generar economías de escala y aprovechar al máximo la infraestructura estadística disponible para viabilizar los programas de encuestas en el largo plazo. Para ello, se enfatizó la importancia de que los INE definan un plan estratégico de mediano y largo plazo que identifique las necesidades futuras de información. Este tipo de ejercicios permite establecer un programa de trabajo que especifique el número de encuestas a realizar, así como los temas que se abordarán, la cobertura geográfica y los dominios analíticos para los cuales se desea obtener información. Estos insumos ayudarán a dar respuesta, entre otras, a preguntas como las siguientes: ¿de qué tamaño deberán definirse las unidades de marco?, ¿cuántas etapas se requiere en el diseño de la muestra para llegar a la unidad de análisis?, y ¿cuántas réplicas es posible reproducir por cada unidad primaria de muestreo?.

Finalmente, se destacó la importancia de que exista una correspondencia adecuada entre los marcos metodológico y conceptual utilizados en los ejercicios mencionados anteriormente, así como la necesidad de uniformizar los clasificadores con los que se codifica la información sobre las características sociodemográficas de los miembros del hogar, tales como escolaridad, ocupación y rama de actividad.

TEMA 2. Diseño y construcción de los marcos de muestreo

Continuando con las ideas expuestas en el tema anterior, la discusión en esta parte del taller tuvo como tema central el uso de marcos maestros de muestreo, entendidos como una herramienta de gran

utilidad para instrumentar los procesos de diseño y selección de muestras maestras, que puedan ser utilizadas en las diversas encuestas contenidas en un plan estratégico de mediano plazo.

Una de las mayores ventajas de emplear un marco maestro de muestreo es que el costo de su elaboración se reparte entre todas las encuestas que lo utilicen, generando un ahorro de recursos importante. Otro atractivo de esta herramienta estadística es la posibilidad de seleccionar unidades de marco de "*reserva*", que permitan atender demandas de información no incorporadas en los programas de trabajo.

La reciente ronda de censos de población y vivienda del 2000 constituye una coyuntura apropiada para iniciar los trabajos de diseño y elaboración de un marco maestro de muestreo. A su vez, este proceso debiera permitir generar –incorporando los métodos más actuales de sistemas de información geográfica (SIG)– cartografía digitalizada y georeferenciada. Varios participantes argumentaron que esta última característica no es indispensable, ya que la ejecución del trabajo de campo generalmente no requiere una métrica geográfica tan precisa para identificar a las unidades de marco.

La elaboración de un marco maestro de muestreo supone, como primer paso, definir el número y tipo de encuestas a realizar, estableciendo sus órdenes de prioridad y fechas probables de ejecución, así como el número de levantamientos previstos a lo largo del ciclo de vida del marco. Para ello, uno de los mayores desafíos que enfrentan los Institutos Nacionales de Estadística consiste en identificar las necesidades futuras de información de sus usuarios.

Se planteó que el desarrollo inicial del marco maestro no requiere de una gran inversión de recursos, puesto que el censo de población y vivienda ya habría generado el material cartográfico que servirá de base para este propósito, y que la mayor parte del trabajo de conceptualización y diseño del marco es una tarea principalmente de oficina. Empero, resulta crucial disponer de cartografía censal actualizada, que defina con claridad límites identificables en campo para las distintas unidades de marco, y que incorpore algún tipo de información para facilitar el control del tamaño de las unidades de marco y sus probabilidades de selección.

La definición e identificación de las unidades de marco, de sus tamaños y de sus límites son aspectos de gran relevancia en la etapa de diseño y construcción del marco de muestreo. Asimismo, el asumir decisiones sobre los puntos de contacto entre las distintas encuestas, determinar la vigencia del marco y definir los plazos y métodos para su actualización son algunas de las decisiones más trascendentes a encarar en su diseño. Adicionalmente, se plantearon inquietudes en torno al momento más apropiado para construir un nuevo marco, la forma de minimizar el efecto sobre los estimadores cuando se realiza un cambio de marco, y la conveniencia de que coexistan dos o más marcos maestros de muestreo.

Una preocupación compartida por varios participantes tuvo que ver con la búsqueda de un mecanismo eficaz y de bajo costo para mantener actualizado al marco maestro, ya que los países suelen no disponer de los recursos necesarios para registrar de manera continua las variaciones que se producen en las viviendas. Sumado a este desafío, se planteó la posibilidad de que los mecanismos de actualización del marco permitan generar información sobre la evolución de la dinámica poblacional, particularmente en términos del volumen de personas. En este sentido, lo deseable sería que los INE efectúen operativos de campo a gran escala que permitan registrar los cambios dentro y fuera de las unidades de marco y su efecto en los estimadores muestrales; es decir, que se implementen procesos permanentes de actualización del marco de muestreo.

En la actualidad, y ante la falta de mejores opciones, los INE utilizan proyecciones para estimar los totales de población a partir de las encuestas de hogares. No obstante, las hipótesis con las que se proyecta la población no siempre consideran las diferencias entre regiones u otras subdivisiones geográficas menores, lo cual impide identificar las diferencias de comportamiento que se presentan en contextos con una dinámica demográfica más intensa. A su vez, a juicio de los participantes, la

actualización permanente del marco de muestreo permitiría estimar volúmenes de población con una precisión mayor a la que entregan las proyecciones.

TEMA 3. Aspectos estadísticos asociados a la construcción del marco de muestreo

Una vez que el marco de muestreo ha sido preparado para la selección de las viviendas, procede encarar los aspectos del diseño estadístico de la muestra. Para ello, un factor a considerar es la identificación de los dominios de estudio y grupos poblacionales de interés para la investigación. Por ejemplo, es común que la necesidad de captar información sobre las diferencias en contextos geográficos lleve a conformar grupos de observaciones que respondan a la división administrativa del territorio nacional, o al menos a la desagregación urbano-rural.

Dado que la noción de estratificación es intrínseca a cualquier investigación por muestreo, resulta ventajoso hacer un uso adecuado de dicho proceso. Una de las ventajas de la estratificación es que permite agrupar observaciones con características similares, con el objeto de dar respuesta a las necesidades específicas planteadas por los usuarios de la información. Por otra parte, desde el punto de vista estadístico también surge la necesidad de clasificar a las observaciones con algún tipo de criterio que permita identificar grupos con características similares.

Hubo consenso en que cualquier procedimiento de estratificación debe orientarse a mejorar la precisión de los estimadores y aumentar la eficiencia del diseño al momento de asignar la muestra, lo cual conlleva la aplicación de algún algoritmo matemático que permita reducir la varianza al interior de las agrupaciones formadas y maximizarla entre los distintos grupos. Este ejercicio se conoce con el nombre de "estratificación óptima".

Los participantes señalaron que, pese a la necesidad de llevar a cabo este tipo de prácticas para mejorar la precisión de los estimadores muestrales, su aplicación por parte de los INE no es muy frecuente. Se indicó que la gran variedad de temas que se estudian en las encuestas, así como los múltiples objetivos que se persiguen en cada una de ellas, limita la aplicación de métodos estadísticos de estratificación óptima. De hecho, en la literatura sobre el tema se reconoce la existencia de limitaciones de carácter práctico que inhiben la aplicación directa de algoritmos multivariantes en encuestas multipropósitos.

En todo caso, una práctica común durante la construcción del marco de muestreo es la clasificación de las unidades según su nivel de bienestar, conforme a la percepción subjetiva sobre el nivel de desarrollo y urbanización de la zona geográfica en que se encuentran, que se infiere a partir de las características de la infraestructura perdurable que los encuestadores identifiquen. Esta práctica permite efectuar la asignación de la muestra conforme a la noción de nivel socioeconómico, situación que resulta de gran utilidad cuando se trata de una encuesta de presupuestos familiares o de condiciones de vida.

Un tema recurrente durante los debates fue la aparente falta de representatividad de las encuestas en el caso de las familias de altos ingresos. Al respecto, se señaló que el esquema de selección de la muestra asigna probabilidades en función al tamaño de las unidades de marco –primera y segunda etapa–, por lo que las zonas densamente pobladas son seleccionadas con mayor probabilidad. Por lo tanto, si la estructura socioeconómica de un país está caracterizada por la mayor presencia de familias de ingresos medios y bajos, esta situación será reproducida por el diseño de la muestra.

En caso que se requiera, ante esta situación es factible sobrerepresentar en la muestra a las familias de altos ingresos. Sin embargo, esta acción deberá ser reforzada con un trabajo de campo que minimice la pérdida de observaciones por rechazo de los informantes. Finalmente, cabe destacar que este tipo de consideraciones, aunque contribuye a mejorar la calidad de los datos, no evita la subdeclaración deliberada que de sus ingresos hacen los informantes.

Otro de los aspectos discutidos en el taller fue la evaluación de la consistencia de la información generada por encuestas mediante su comparación con otras fuentes de datos. En particular, se aludió a la conciliación de un conjunto básico de variables sociodemográficas de la encuesta con los resultados del

censo de población y vivienda, sobretodo cuando ambas fuentes corresponden a períodos cercanos. En estos casos, se aconsejó utilizar la información censal con criterios de post-estratificación. Por ejemplo, se consideró pertinente ajustar los datos de las encuestas para reflejar los cambios en la estructura por edad y sexo reportada en la información censal. Esta práctica se basa en el supuesto de que el censo refleja mejor la dinámica demográfica del país, y, por tanto, provee puntos de partida adecuados para la conformación de nuevas series de variables sociodemográficas durante los períodos intercensales.

A lo largo del taller, se planteó varias veces la reflexión sobre qué hacer cuando existen nuevas construcciones y cómo incorporarlas en los procesos de actualización del marco de muestreo. Es evidente que los acelerados cambios inducidos por la urbanización de las poblaciones obligan a los muestristas a estar atentos a la dinámica demográfica.

TEMA 4. Aspectos operativos para la construcción del marco de muestreo

El uso de diseños de muestra maestros, cuyo propósito es efectuar selecciones múltiples de viviendas que permitan atender las demandas de información de una o más encuestas periódicas o multirondas, fue uno de los aspectos operativos del marco de muestreo discutidos en el taller.

Entre las principales ventajas de trabajar con un diseño de muestra maestro –cuya lógica de construcción mantiene una correspondencia directa con la del marco maestro– está el disponer de una o más muestras maestras, seleccionadas a partir de criterios uniformes y con el mismo número de etapas. Se señaló que es factible utilizar en todos los casos las mismas unidades de marco, aplicar el mismo procedimiento de cálculo para determinar las probabilidades de selección y usar el mismo marco maestro de muestreo. No obstante, la eficiencia en la organización y administración de las encuestas puede ser mejorada empleando variaciones a este sistema.

Por otra parte, se insistió en la importancia de tomar en cuenta los distintos objetivos y coberturas geográficas de las encuestas, así como sus dominios de estudio, al momento de aplicar un procedimiento de estratificación para agrupar las distintas unidades de marco.

Una de las interrogantes planteadas respecto a este tema fue si lo que se debe actualizar es la muestra maestra o el marco maestro que dio origen a esa muestra –cabe destacar que ambas posibilidades son conceptualmente distintas, y que la respuesta incide en el procedimiento de trabajo a emplear y en el monto de recursos requerido. Se señaló que, en principio, lo más apropiado parece ser mantener actualizado el marco maestro utilizado para la selección de viviendas, sin dejar de lado la preocupación de que al hacerlo se asume que el crecimiento de las unidades de vivienda fuera del marco sigue el mismo patrón que el que se está observando. En este sentido, se recomendó que la actualización se efectúe como una actividad previa a la realización de las encuestas, ya que de esta manera los cambios observados en el terreno pueden incorporarse en los procesos de estimación, generando un efecto favorable en la calidad y precisión de la información.

Por otra parte, se resaltó que los procesos de actualización del marco pueden significar un porcentaje muy importante del costo total de una encuesta. Sin embargo, el disponer de un marco maestro de muestreo –así como de una muestra maestra– constituye una inversión necesaria para garantizar la calidad de los resultados de las encuestas. Al respecto, cabe insistir en que, aunque se emplee el mejor diseño de muestra y se defina el método de estimación más robusto, la calidad de la información y el cumplimiento de los objetivos de la investigación están en riesgo si no se tiene un marco de muestreo adecuado.

Los participantes de los distintos países informaron acerca de los procedimientos que utilizan sus oficinas para realizar los recorridos dentro de las unidades seleccionadas, así como aquellos que se aplican para elaborar y actualizar los listados de viviendas y la identificación de hogares que no han sido detectados en la cartografía censal. Quedó en evidencia que la mayoría de los países destina recursos

escasos o nulos a la actualización periódica del marco, hecho que hace necesario recurrir a las proyecciones de población –exógenas a la encuesta– para generar estimaciones de los totales de personas.

En cuanto al tratamiento que los países hacen de las viviendas adicionales y omitidas, fue posible constatar que no es común el que se efectúen acciones específicas para resolver este problema, que forma parte de las limitaciones habituales que afectan a los marcos de muestreo.

TEMA 5. Sistemas de Información para la selección de la muestra

La automatización de los procesos de trabajo se destacó como una preocupación permanente por parte de los participantes, que reconocieron en su aplicación a procesos como la selección de las unidades de marco, la actualización de los listados y la selección de las viviendas la posibilidad de desempeñar estas tareas de manera más eficiente y liberar tiempo y recursos para dedicarlos a otras actividades.

Uno de los puntos discutidos fue la posibilidad de diseñar un sistema de información que combine de manera adecuada el manejo de la cartografía y los procesos estadísticos asociados al diseño de la muestra. Específicamente, la idea es que las actualizaciones identificadas durante el trabajo de recolección de información en campo puedan ser incorporadas a una base de datos, estructurada como un Sistema de Información Geográfica, y que a partir de ello se realicen automáticamente los ajustes necesarios a los factores de expansión. Algunos de los países participantes en el taller están realizando actualmente esfuerzos en esta dirección; en particular, la delegación del Perú expuso su avanzada experiencia al respecto, y se destacó que en Venezuela y México están en marcha proyectos con objetivos similares.

Las opiniones con respecto a esta iniciativa fueron muy favorables, considerando que la selección de las viviendas, el control del trabajo de campo y el ajuste posterior de las probabilidades de selección son actividades mecánicas que, realizadas de forma manual, están sujetas a un alto grado de error y consumen un tiempo excesivo. Adicionalmente, la automatización de procesos contribuye a elevar su legitimidad y transparencia, y permite ampliar la memoria institucional respecto a las decisiones tomadas durante la ejecución del trabajo de campo, que tienen un importante efecto sobre los resultados de las encuestas.

Como beneficios derivados de los sistemas de información, también se señalaron la reducción en uso de papel y ahorro del espacio físico requerido para guardar los listados impresos, y la posibilidad de atender demandas de información especializada referida a los residentes y negocios comerciales ubicados en las áreas de listado.

LOS CENSOS DE POBLACIÓN Y VIVIENDA Y LOS MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES

FERNANDO MEDINA H.
CEPAL

Programa para el Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe
MECOVI

9º Taller Regional
Diseño y Construcción de Marcos de Muestreo para las Encuestas de Hogares

Los Censos de Población y Vivienda y los Marcos de Muestreo para las Encuestas de Hogares

Fernando Medina H.
CEPAL
Junio de 2002
Lima, Perú

¿Porqué es importante coordinar las actividades del Censo de Población y Vivienda con los programas de encuestas?

- La realización de los Censos de Población y Vivienda, representa, para muchos países, la única oportunidad para actualizar y renovar el material cartográfico e informativo que se utiliza para planear y ejecutar las encuestas en hogares
- El material y la información que generan los censos de población y vivienda, se consideran insumos fundamentales para enfrentar con éxito el diseño y desarrollo de un programa de Encuestas en Hogares: **renovación del marco de muestreo**
- La adecuada coordinación de estas actividades, genera sinergias positivas que se traducen en economías de escala y ahorros de recursos

¿Qué comparten?

- recursos
- definiciones
- contenido temático: variables sociodemográficas
- materiales y marcos
- clasificadores

¿Cuáles son las similitudes que existen entre ambos eventos?

- en parte, tienen en común las mismas **unidades de selección: viviendas**
- es común para ambos las **unidades de observación: las familias**
- comparten el interés por la misma **población objetivo: personas**
- actúan sobre la misma **estructura geográfica: el territorio nacional**

¿Qué define la relación entre ambos eventos?

- el **marco de muestreo** que sirve de referencia para la ejecución de las encuestas en hogares, es un **subconjunto** del marco de edificaciones construido a propósito del Censo de Población y Vivienda

¿Cuáles son las razones técnicas y administrativas de esta relación?

el Censo genera el marco de muestreo de viviendas

- reducción de costos
- buscar coherencia conceptual
- identificación de variables comúnes
- incoprorar las variables necesarias para la estratificación del marco de muestreo
- unidades administrativas

continuación

- unidades de marco, mapas, límites en campo, listados de viviendas
- archivos magnéticos
- catálogos
- estrategias para la ejecución del trabajo de campo
- logística operativa
- vehículos, materiales y otro tipo de insumos

¿ En qué momento inicia la coordinación de los trabajos para la elaboración del marco de muestreo?

- Desde el momento en que se está planeando el **pre-censo** es necesario coordinar los trabajos, de tal suerte que se generen las mayores sinergias para la elaboración del marco de muestreo
- No debemos olvidar, que el censo de población también requiere disponer de un buen marco que incorpore al conjunto de las edificaciones que conforman el país:viviendas comercios, escuelas, instituciones, etc.
- Se incluyen viviendas particulares y colectivas, negocios, instalaciones deportivas y para recreación, etc.

continuación

- durante el pre-censo y la ejecución del censo, se debe pensar en la **estructura que tendrá el marco de muestreo: programa de encuestas de hogares**
- identificar la información que pueden ser útiles para: definir las unidades de marco, las definiciones y variables que se utilizarán en la etapa de estratificación
- características que debe tener el material auxiliar (cartografía, listados de edificaciones, bases de datos, etc)

¿Existe armonía entre los materiales del censo y los que requiere el marco de muestreo?

es habitual que se deban hacer adaptaciones

- los censos de población basan su estructura operativa en divisiones administrativas oficiales, las cuales no siempre se corresponden con las que se requieren para la elaboración del marco de muestreo
- normalmente, existe un desfase entre la fecha en que se elabora la cartografía censal, y la fecha en que se diseña y construye el marco de muestreo: **actualización**
- los segmentos o áreas censales no siempre se corresponden con las unidades de marco: **no importa**

¿Cuáles son los inconvenientes más habituales?

- **desactualización del marco**: diferencias entre la realidad y los mapas y planos preparados para la ejecución del censo
- es necesario redefinir el concepto de unidades de marco y su tamaño óptimo, para fines de la elaboración del marco de muestreo conforme a las necesidades del programa de encuestas: **marco maestro (sistema integrado de encuestas de hogares) vs marcos individuales (encuestas integradas)**
- la definición y conformación de las unidades de marco, significa partir o fusionar segmentos o áreas censales, conforme a las necesidades especificadas en el **diseño de muestra**

¿En qué momento finaliza la relación entre el marco de hogares y el marco del censo?

- toda vez que se definieron las unidades de marco (primera etapa: **marco de áreas**), su tamaño, sus límites, su identificación, su clasificación por estrato, de alguna manera se pierde la relación biunívoca entre el marco del censo de población y el diseñado para planear y ejecutar las encuestas de hogares
- a partir de ese momento, las prioridades del área de muestreo y de marcos, se deben orientar a mantener actualizado el marco de referencia para la ejecución de las encuestas

¿De qué manera se mantiene la relación entre ambos instrumentos?

- los resultados del censo de población y vivienda, se utilizan como insumos informativos para que los demográfos realicen proyecciones de población
- en este sentido, las proyecciones de población, basadas en cifras censales, en muchos países se mantienen como referente para efectuar estimaciones a partir de las encuestas de hogares

¿Cómo se está aprovechando la información censal más allá de su divulgación?

- una práctica que ha cobrado vigencia en América Latina y el Caribe, es el uso de la información de las encuestas para parametrizar los datos de hogares del Censo de Población, con el objetivo de efectuar estudios de pobreza y bienestar
- algunas precauciones:
  - las fechas de las encuestas y la correspondiente al censo de población
  - los cambios en la estructura por edad y sexo en algunas de las localidades del país
  - migraciones, nuevas construcciones, destrucciones, etc.

¿Cuál es la información básica que se requiere para elaborar el marco de muestreo?

- **áreas censales** bien definidas y delimitadas
- **medidas de tamaño** de cada una de ellas: total de viviendas, personas –por edad y sexo-
- **cartografía** básica digitalizada: no es necesario georeferenciada
- **definiciones** censales básicas
- **divisiones administrativas** censal y su jerarquía
- **datos básicos**: promedio de hogares por vivienda, tamaño medio del hogar, etc.

¿Preguntas importantes para la reflexión?

- ¿realmente existe una **vinculación** entre la planeación del censo y el diseño del marco de muestreo?
- ¿**qué se está haciendo** en los países en esta materia?
- ¿**participan los estadísticos** de encuestas y la sección de marcos estadísticos en el diseño conceptual y metodológico del censo?
- ¿**participan los estadísticos** de encuestas en la definición del cuestionario censal?

ASPECTOS ESTADÍSTICOS ASOCIADOS A LA CONSTRUCCIÓN DEL MARCO DE MUESTREO

FERNANDO MEDINA H.
CEPAL

Programa para el Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe
MECOVI

9º Taller Regional
Diseño y Construcción de Marcos de Muestreo para las Encuestas de Hogares

Aspectos estadísticos asociados a la construcción del marco de muestreo

Fernando Medina H.
CEPAL
Junio de 2002
Lima, Perú

La instrumentación del diseño muestral y la construcción del marco de muestreo

- el **diseño y la elaboración del marco** de muestreo es una de las etapas asociadas a la instrumentación y ejecución del diseño muestral
- la **definición** de las unidades de marco, las de selección y el número de etapas en que éstas serán seleccionadas, responden a decisiones que han sido plasmadas en el diseño de muestra
- de alguna manera la **ejecución del diseño** da inicio con la elaboración del marco muestral, conforme al esquema de selección definido

¿Qué es estratificar?

- consiste en **dividir** a la población objeto de estudio en dos o más grupos con el objetivo de seleccionar muestras independientes en cada uno de ellos
- se forman **grupos de observaciones** con características homogéneas, pero no necesariamente a partir de la aplicación de criterios estadísticos
- la clasificación de las observaciones en el contexto regional o urbano-rural, debe ser entendido como un ejercicio de estratificación

¿Porqué es necesario estratificar?

- desde el **punto de vista estadístico**, la estratificación se justifica como una manera de **reducir la varianza** de los estimadores muestrales:
  - La **varianza disminuye** según el grado de homogeneidad que exista entre los elementos al interior de los estratos
  - la varianza puede disminuir asignado mayor muestra en estratos que presentan mayor variación o menor costo
- Se **estratifica**, con el objeto de utilizar los estratos como dominios de estudio, y es posible utilizar distintos métodos de selección al interior de cada uno de ellos

continuación

- también se aplica cuando los objetivos de la investigación demanda **estimadores con precisiones distintas por estrato**
- por conveniencia práctica y administrativa

¿Todas las estratificaciones son estadísticas?

- **No.** la estratificación típica para clasificar las unidades primarias de muestreo es la geográfica: urbano-rural, región, provincia, tamaño de la localidad
- es común que las **estratificaciones administrativas** puedan requerir más de una clasificación de las observaciones: regional y urbano-rural
- también es una práctica habitual que las divisiones geográficas se utilicen como dominios de estudio independientes
- a su vez, estos dominios pueden ser subdivididos en estratos

Estratos especiales

- terrenos baldíos destinados a la construcción de fraccionamientos habitacionales
- campamentos: tiendas de camapaña, trailers, damnificados
- ciudades deportivas
- poblaciones especiales
- es habitual que las unidades de penúltima etapa se estratifiquen por nivel socieconómico, lo cual implica diseñar criterios de campo para clasificar a cada unidad en un estrato

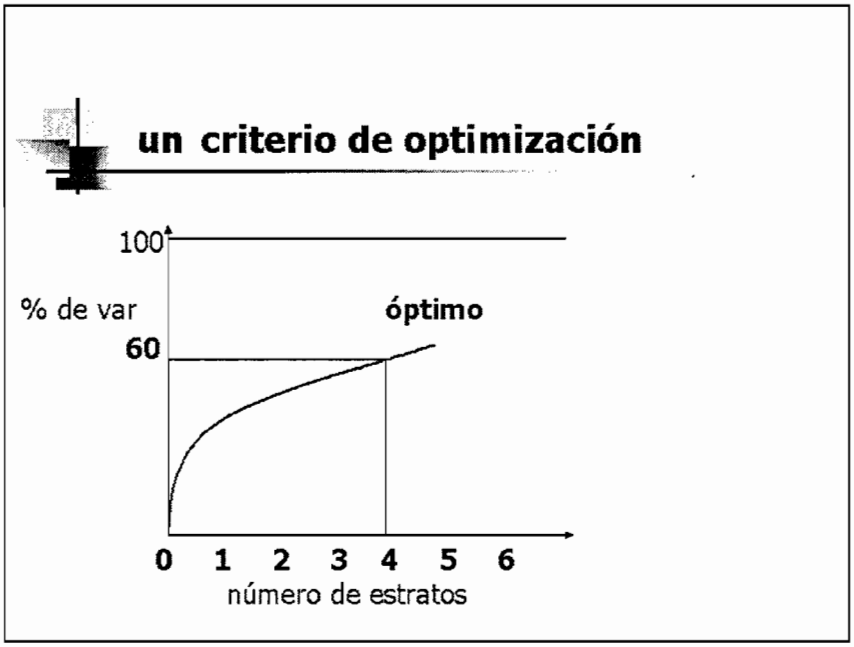
Métodos estadísticos de estratificación

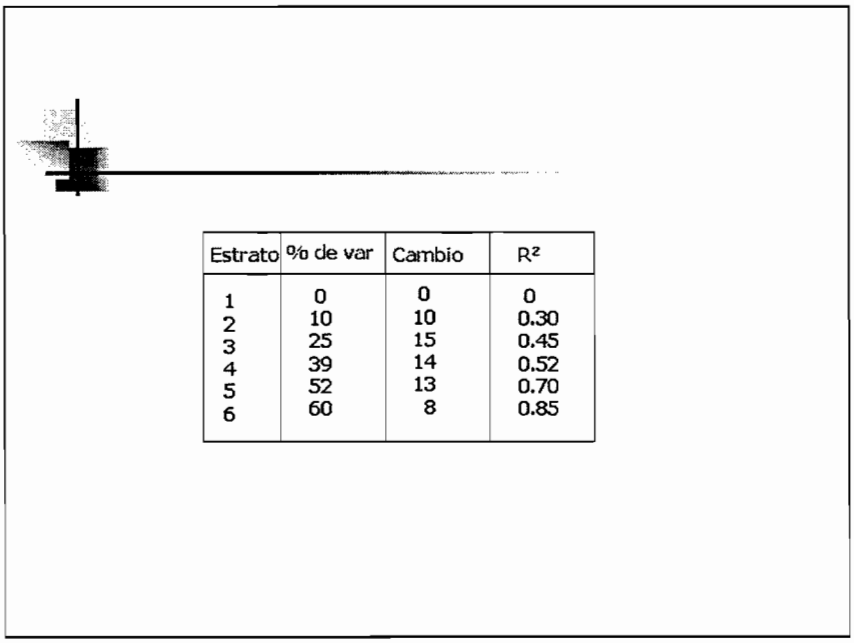
- las **estratificaciones estadísticas** requieren la aplicación de algoritmos de optimización
- se parte del interés de **optimizar una función criterio**
- se busca generar **grupos homogéneos en su interior** y la mayor **heterogeneidad entre grupos**
- para el caso univariado se tienen los algortimos propuestos por Dalenius y Hodge (1959) y (1963)

- en el caso multiparamétrico existen diversos algoritmos (Hartigan, 1975)
- entre otros, el algoritmo de las **k-medias** (MacQueen, 1967) que utiliza criterios de partición en la formación de los conglomerados es ampliamente utilizado
- en **Jarque (1981)** se presenta una solución al problema de estratificación óptima en muestreo multivariado, en el cual se combina el método de Dalenius con un algoritmo propuesto por Sparks(1973)

Preguntas importantes

- ¿cuál es el **número óptimo de estratos**?
- ¿cómo se indentifican sus **límites**?
- ¿cómo se conoce el **porcentaje de varianza explicada**?
- ¿cuál es el **criterio de distancia** para agrupar estratos?
- ¿cuál es el criterio para identificar la **estratificación óptima**?





Estrato	% de var	Cambio	R^2
1	0	0	0
2	10	10	0.30
3	25	15	0.45
4	39	14	0.52
5	52	13	0.70
6	60	8	0.85

Términos comúnes

- **estratificación primaria** corresponde a las clasificaciones de las UPM´s
- por su parte, el témino **subclase** se refiere a una colección de unidades con ciertas caracteríticas de interés
- es común que posterior a las estratificaciones geográficas se aplique una estratificación estadística con una o más variables de interés
- una estratificación que conviene a los intereses de la encuesta y que tiene como objetivo **minimizar la varianza de muestreo**

¿Qué objetivo se persigue al asignar la muestra en los estratos?

- el **objetivo** es encontrar un esquema de asignación que genere **estimadores insesgados** y sea eficiente en términos de varianza
- **la optimización** en la asignación de la muestra, **depende** del tamaño de los estratos y además de la varianza de la variable de estudio al interior del estrato y entre estratos

¿Cómo se puede asignar la muesrta en los estratos?

- **Asignación Proporcional**: es el más simple de los métodos de asignación y es muy común su uso en la práctica
- Presupone que **se conoce el tamaño del estrato**, debido a que la fracción de muestro es constante en cada estrato: $\mathbf{n_h/N_h}$
- el número de elementos por estrato es:

$$\mathbf{n_{hpro} = n \times N_h / n = n \times W_h}$$

- Mediante este método de asignación **se garantiza la misma proporción de muestra en todos los estratos**
- como la fracción de muestreo es constante en cada estrato la **probabilidad de selección** de cualquier elemento es n/N

- este esquema de asignación resulta equivalente a un esquema aleatorio simple, lo cual favorece el procedimiento de estimación

$$\hat{t}_{estr} = \hat{t} = N \sum_{h=1}^{H} \sum_{k=1}^{n_h} \frac{y_{hk}}{n}$$

- las medias al interior de cada estrato no requieren ser calculadas
- por esta razón, la asignación proporcional cumple con la propiedad de asignar una **muestra autoponderada**

Genera los estimadores más eficientes en el muestreo estratificado

- **Asignación de Neyman**: requiere **conocer la desviación estándar** de la variable de interés al interior de cada estrato S_h
- el total de elementos por estrato se obtiene mediante:

$$n_{hopt} = n \frac{N_h S_h}{\sum_{h=1}^{H} N_h S_h}$$

- en la práctica, es difícil conocer el valor de S_h, el cual se debe aproximar a partir de la experiencia o encuestas previas
- un estrato grande o con mucha variación interna debe tener más elementos en muestra

¿En qué situaciones se sugiere aplicar?

- **"Power Allocation"**: cuando existen **muchos estratos con pocos elementos** cada uno y además se requieren estimaciones precisas en cada uno de ellos
- el número de asignaciones por estrato para estimaciones eficientes de totales está dado por:

$$n_{hpower} = n \frac{(T_{zh})^a cv(y_h)}{\sum_{h=1}^{H} (T_{zh})^a cv(y_h)}$$

- T_{zh} es el total en el estrato de una variable auxiliar z y $cv(y_h)$ el coeficiente de variación de y en el estrato h.

- La constante **a** es conocida como la **potencia** de la asignación y en la **práctica asume** valore de ½ o 1/3
- Esta método puede ser visto como una **asignación de compromiso** entre la de Neyman y aquella que le asigna aproximadamente precisión constante a todos los estratos

Ejemplo

- tenemos una población dividida en **dos estratos**: urbano y rural
- existen **32 municipios** de los cuales 7 son urbanos y 25 rurales
- se requiere seleccionar una **muestra aleatoria de 8 municipios**
- se aplicarán los **tres métodos de asignación** de la muestra
- la variable total de hogares en 1990 se utiliza como información auxiliar para la asignación potencia

Información relevante

Parámetro	Variable	Estrato 1	Estrato 2	Total
Media	x_1	1,146	283	472
Total		8,022	7,076	15,098
Desv. Est.		1,318	331	743
Coef. Var.		1.150	1.170	1.572
Total	**hogares en 90**	**49,842**	**41,911**	**91,753**

Resultados

- **Asignación proporcional:** $n_{hpro} = n\frac{N_h}{N} \rightarrow \left\{ \begin{array}{l} n_1 = 8X\frac{7}{32} = 1.75 \\ n_2 = 8X\frac{25}{32} = 6.25 \end{array} \right\}$

- **Asignación Optima:** $n_{hopt} = n\frac{N_h S_h}{\sum_{h=1}^{H} N_h S_h} \rightarrow \left\{ \begin{array}{l} n_1 = (8)\frac{9226}{9226+8275} = 4.22 \\ n_2 = (8)\frac{9226}{9226+8275} = 3.78 \end{array} \right\}$

- **"Asignación Potencia":** $n_{ha-1} = n\frac{T_{zh}cv(y_h)}{\sum_{h=1}^{H} T_{zh}cv(y_h)} \rightarrow \left\{ n_1 = (8)\frac{49842x1.15}{57318+49036} = 4.31 \right.$

$$n_2 = (8)\frac{41911x1.17}{57318+49036} = 3.69$$

continuación

$$n_{h,a=\frac{1}{2}} = n\frac{\sqrt{T_{zh}}cv(y_h)}{\sum_{h=1}^{H}\sqrt{T_{zh}}cv(y_h)} \rightarrow \left\{ n_1 = 8\frac{\sqrt{49842}x1.15}{257+240} = 4.13 \right.$$

$$n_2 = 8\frac{\sqrt{41991}x1.17}{257+240} = 3.86$$

$$n_{h,a=0} = n\frac{cv(y_h)}{\sum_{h=1}^{H}cv(y_h)} \rightarrow \{ n_1 = 8\frac{1.15}{1.15+1.17} = 3.97$$

$$n_2 = 8\frac{1.17}{1.15+1.17} = 4.03$$

comparaciones

Asignación	Elementos por estrato	
	n_1	n_2
Proporcional	2	6
Optima	4	4
Potencia	4	4

La mayor eficiencia de la asignación óptima y la potencia respecto a la proporcional, se advierte al calcular el efecto de diseño $deff_{(opt,pot)}$=0.44 y $deff_{pro}$=0.74

Etapas para el diseño de la muestra

- definir con claridad los **objetivos** de la investigación
- definir las **unidades de observación y análisis**
- definir los **dominios análiticos** de interés
- definir el **número de etapas** en que se seleccionará la muestra
- identificar las **unidades a seleccionar** por etapa
- decidir acerca del **método de selección** en cada etapa
- fijar el **tamaño de muestra: variables de interés**

- fijar la **tasa de muestreo** por etapa de selección
- definir la **precisión** requerida para las distintas variables de interés
- **formar estratos**
- **formación de conglomerados**

Diseños multietápicos

- **todos los diseños** de las encuestas de hogares **son multietápicos**
- el costo estadístico que tienen los **diseños multietápicos** es que **incrementan el error de muestreo,** comparado con un diseño aleatorio simple
- el **incremento en la varianza** depende de la homogeneidad de los elementos respecto a la variable de interés en las primeras etapas, del esquema de selección y del número de selecciones en cada etapa
- la **pérdida de eficiencia** puede aumentar, si el número de selecciones por conglomerado aumenta o si las unidades de selección están dispuestas en unidades compactas

ventajas

- el trabajar con diseños multietápicos **concentra el trabajo de campo y reduce los costos operativos**
- **abarata el costo** de diseño y mantención del marco de muestreo
- la primera etapa de un diseño multiétápico es la que mayor relación guarda con la información y cartografía censal
- las etapas posteriores son independientes

La estratificación en los diseños multietápicos

- es muy común que en los diseños multietápicos se apliquen criterios de estratificación, lo cual habitualmente se traduce en mayor precisión de los estimadores
- en la práctica, es común que existan criterios administrativos y de otro tipo para estratificar las unidades de primera etapa (UPM's)
- la estratificación en los diseños polietápicos permite un **mayor control** en la asignación y selección de la muestra

Ventajas de estratificar en los diseños multietápicos

- el costo de la estratificación delas UPM's tiende a ser bajo en comparación con sus beneficios
- se logra mayor control en la selección y asignación de la muestra

Los diseños de tres o más etapas

- las etapas de selección en un diseño de muestra está en función a dos factores de costo: **listados de viviendas y traslado de los encuestadores**
- **cuando se vuelve muy costosa** la, subdivisión de unidades de primera etapa de gran tamaño, lo aconsejable es definir unidades de segunda etapa y muestrear sobre ellas
- el **número de etapas** para la selección de la muestra, se define en función de los costos y las ventajas operativas que se tengan para la ejecución del diseño muestral

Características de las unidades primarias de muestreo (UPM'S)

- el número de **unidades primarias** debe ser grande, por lo menos 5 ó 6 veces mayor al número de selecciones que deseamos efectuar
- deben tener **límites bien identificables y medidas de tamaño robustas**
- se busca que **en promedio sean del mismo tamaño,** aunque en la práctica se acepta cierta dispersión en torno a su media
- que sean fácilmente **identificables en el campo**

¿Cómo se selecciona la muestra?

- es común que en las encuestas de hogares las unidades de pirmera etapa se selecciones sietmáticamente con **probabilidad proporcional a su tamaño (ppt)**
- la **medida de tamaño** comunmente utilizada es el total de personas o el total de hogares
- este esquema representa una manera conveniente para seleccionar la muestra **cuando existen variaciones importantes en el tamaño de las unidades de selección**

¿Cómo se selecciona la muestra?

- en los diseños de tres o más etapas es necesarios subdividir las unidades de **primera y segunda etapa,** por lo que es comú que la selección en las primeras etapas se controle a partir de muestreo con ppt
- es común que las unidades últimas de muestreo se seleccionen en forma **sistemática con igual probabilidad**
- en ocasiones, las selecciones de primera etapa se efectúa con **selecciones pareadas,** pero esta tarea requiere de mayor cuidado aunque el cálculo de varianzas es más sencillo

El nuevo marco muestral y su relación con el censo

- existe una **relación** de los instrumentos censales con el marco de muestreo
- no obstante, en ocasiones esta se puede diluir rápidamente
- es posible que las áreas censales coincidan con las unidades de primera etapa del marco de muestreo, pero también existen situaciones en que no tiene porque existir una relación biunívoca
- las áreas de enumeración censal se puede dividir, agrupar e incluso expander en tamaño, en caso de que esto sea necesario para los fines del diseño de muestra

El nuevo marco muestral y su relación con el censo

- sin embargo, es recomendable que se tenga precaución de establecer un registro riguroso de los procedimientos aplicados y los cambios efectuados, en el entendido de que en algún momento se considere necesario reconstruir las áreas censales
- para lograr **armonía** entre las unidades del censo y las del marco de muestreo, se sugiere que los muestristas y los responsables del marco, tengan una participación activa en la definición de los criterios que se aplicarán para la construcción de las unidades censales

¿Qué hacer para construir el marco?

- definir lo que se entenderá por **unidades de marco**
- definir el **tamaño** de las unidades en cada etapa
- **identificarlas** en los mapas: agregarlas, dividirlas
- definir, identificar y trazar sus **límites**
- **estratificarlas**
- **enumerarlas**
- **asignarles** medidas de tamaño

¿Qué hacer para construir el marco?

- **preparación** de **los marcos** de muestreo
- **identificación de las manzanas o áreas de listado** y sus límites
- **selección** de las unidades de marco
- **elaboración** de los **listados** de viviendas
- **formación de segmentos** de viviendas
- **actualización del marco**: crecimiento demográfico, extensión de la mancha urbana, etc.

Algunas preguntas relevantes

- ¿Cómo se trabaja en los países?
- ¿Cuáles son los criterios que se aplican para definir el tamaño de las unidades de marco?
- ¿Se trabaja bajo el concepto de marco maestro?
- ¿Cómo se estratifica?
- ¿Qué algortimo de estratificación se aplica?
- ¿Cuáles son los criterios de estratificación?

Algunas preguntas relevantes

- ¿Quiénes son los encargados de realizar estas tareas?
- ¿Existe comunicación entre las áreas de censos y las de muestreo?
- ¿Cómo se documentan los procesos?
- ¿Cómo se actualiza y mantiene vigente el marco de muestreo?
- ¿Cada cuándo se actualiza en campo el marco?

¿Qué le falta a los diseños de muestra actuales, para satisfacer las necesidades de diseño y evaluación de políticas?

- Número de estratos
- Dominios de estudio e inferencia estadística
- Tamaño de muestra
- Estratos especiales ¿Cómo muestrear?
- ¿Cómo ponderar?
- ¿Cómo ajustar los factores de expansión?
- Exclusión de poblaciones extrañas
- Post-estratificación

Encuestas de Panel

- ¿Para qué?
- Porcentaje óptimo de traslape
- Seguimiento y evaluación de políticas
- Seguimiento de personas u hogares
- Estimadores de cambio
- Traslape de marcos

Estimación y ponderación

- Areas de nuevo crecimiento
- Cambios de marco
- Creación de nuevas unidades administrativas
- Uso de proyecciones de población para expandir resultados

Segmentos compactos Selección sistemática

- Contaminación
- Probabilidades de selección
- Trabajo de campo
- Errores no muestrales
- Asignación de muestra
- Selección controlada
- Comparaciones pareadas

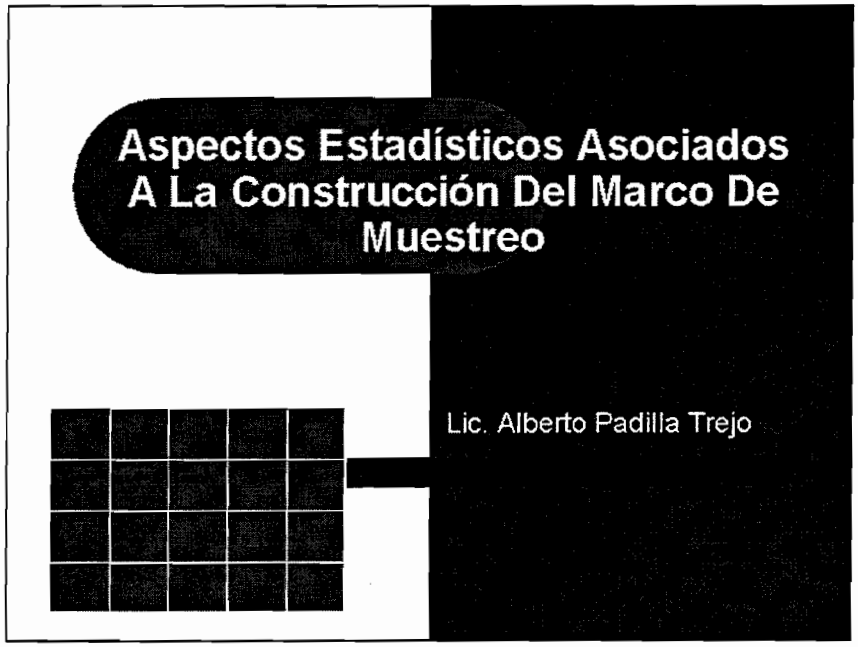
Ejecución del Trabajo de Campo

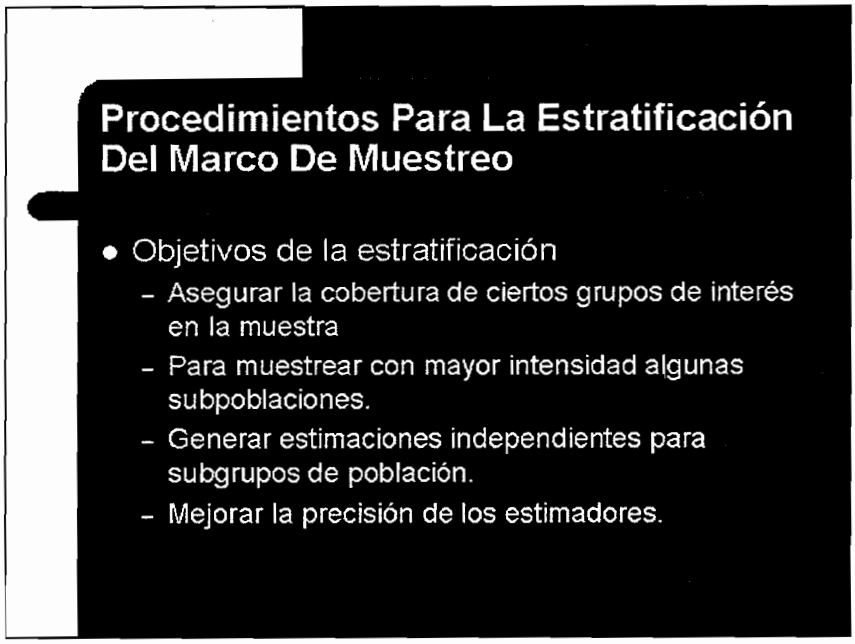
- ¿Criterios a segmentar?
- Tratamiento diferencial: áreas urbanas y rurales, zonas de difícil acceso, áreas de violencia
- Segmentación en zonas densamente pobladas edificios multifamiliares
- Recorridos en campo: viviendas seleccionadas, adicionales, omitidas
- Técnicas y procedimientos para la elaboración de conteos y listados de viviendas
- Partición de manzanas
- Sustitución de viviendas: probabilidades de selección, sesgos de selección, tasa de no respuesta
- Estratificación subjetiva: criterios

ASPECTOS ESTADÍSTICOS ASOCIADOS A LA CONSTRUCCIÓN DEL MARCO DE MUESTREO

ALBERTO PADILLA TREJO

INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA (INEI)
PERÚ


Aspectos Estadísticos Asociados A La Construcción Del Marco De Muestreo
Lic. Alberto Padilla Trejo


Procedimientos Para La Estratificación Del Marco De Muestreo
• Objetivos de la estratificación
- Asegurar la cobertura de ciertos grupos de interés en la muestra
- Para muestrear con mayor intensidad algunas subpoblaciones.
- Generar estimaciones independientes para subgrupos de población.
- Mejorar la precisión de los estimadores.

Procedimientos Para La Estratificación Del Marco De Muestreo

- Criterios para la estratificación explícita
  - Áreas: Urbano – rural
  - Niveles de aglomeración de la población
    - Ej. Poblaciones dispersas, hasta x; hasta nx, hasta mx
  - Importancia administrativa
    - Poblaciones dispersas, resto rural, Ciudades pequeñas, ciudades capitales o cabeceras de distritos, Grandes capitales, Mega ciudades
  - Tipo de actividad predominante

Procedimientos Para La Estratificación Del Marco De Muestreo

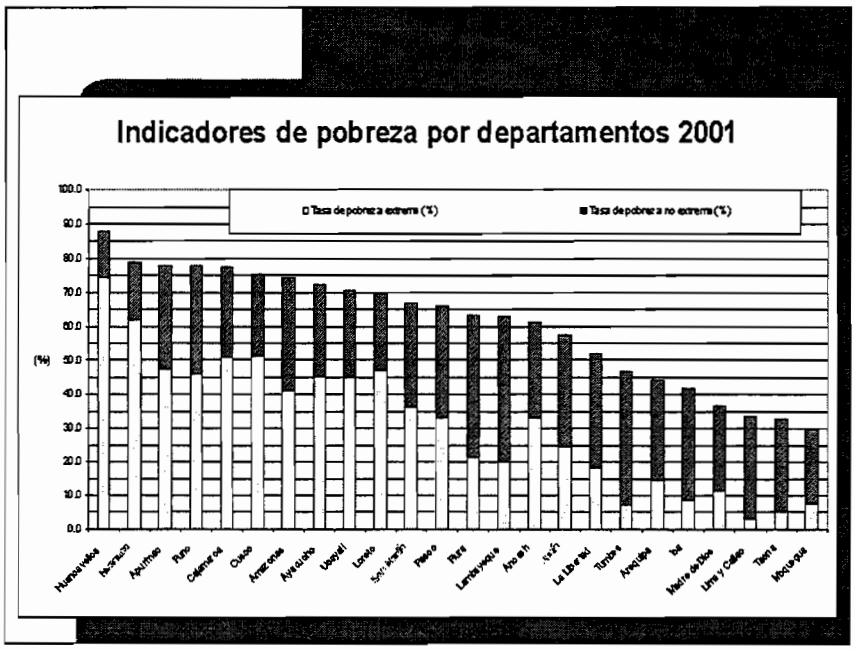
- Criterios para la estratificación explícita
  - Para medición de impacto de políticas (universales o focalizadas):
    - Poblaciones beneficiarias de determinados programas sociales. Tipos o cantidad de programas concurrentes
      - Ejs. Corredores económicos, Mitch.
  - Para medición de acceso a servicios
    - Distancias (radios)
  - Para medición de calidad de servicios
    - Tipos de complejidad del servicio, tipos de establecimientos, tipos de clientes.

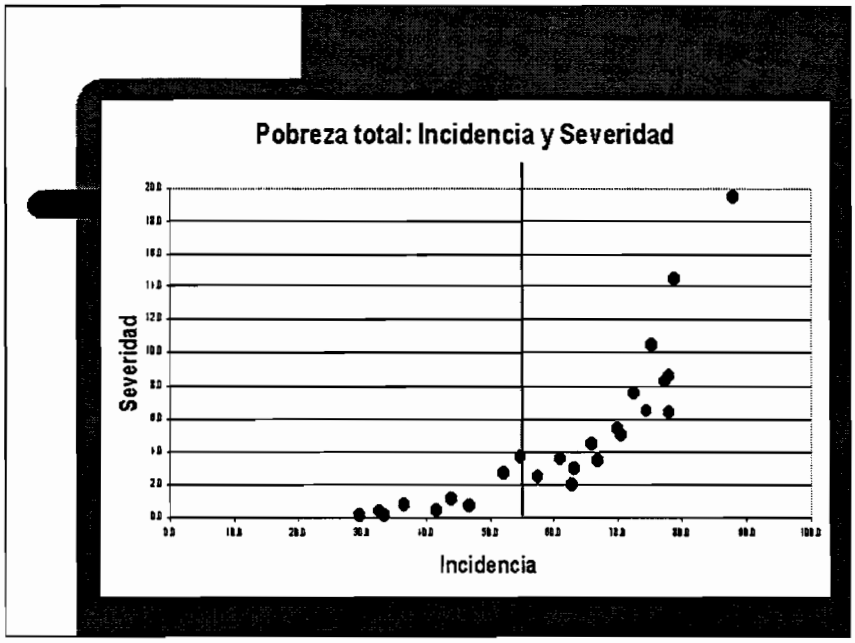
Procedimientos Para La Estratificación Del Marco De Muestreo

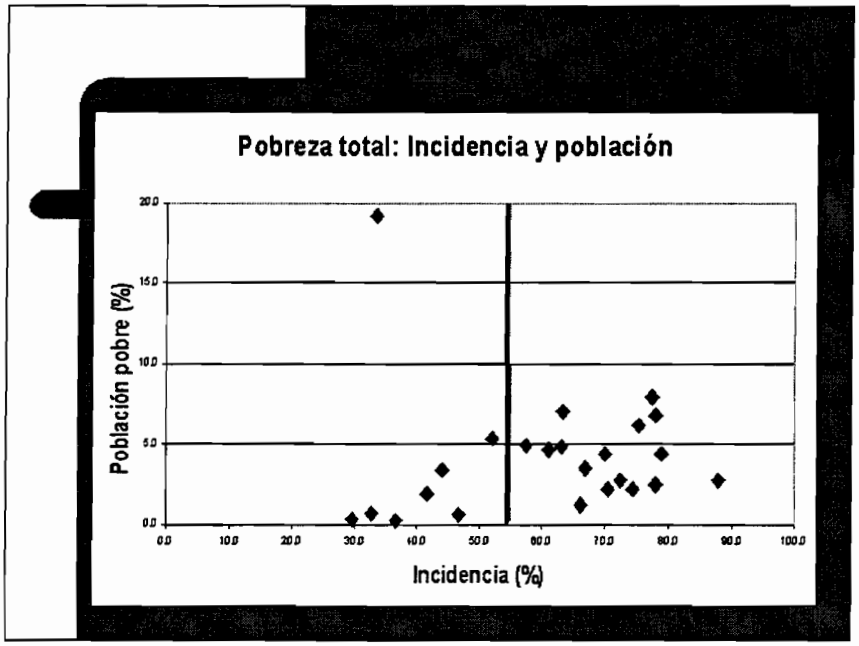
- Criterios para la estratificación implícita
  - Muestreo sistemático sobre la base de un ordenamiento en función de una variable.

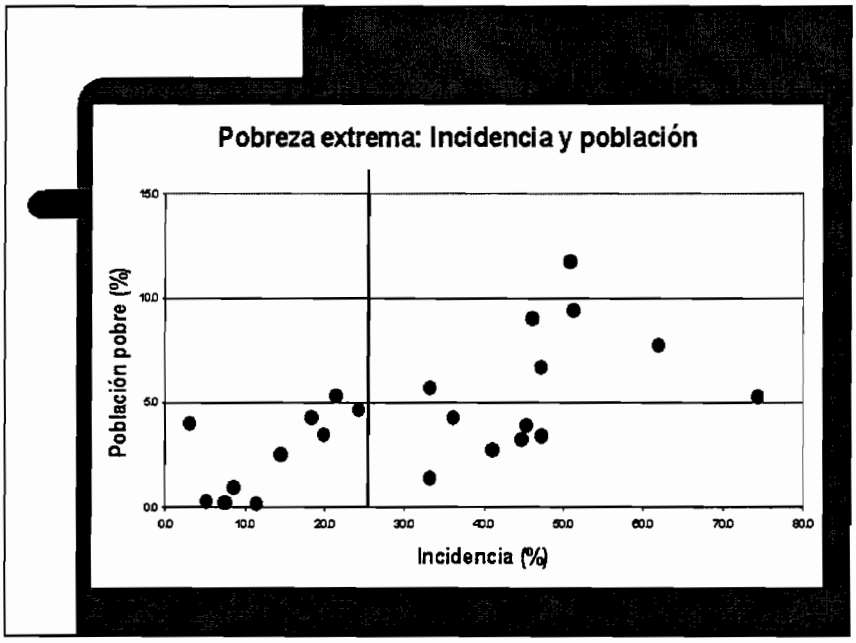
Procedimientos Para La Estratificación Del Marco De Muestreo

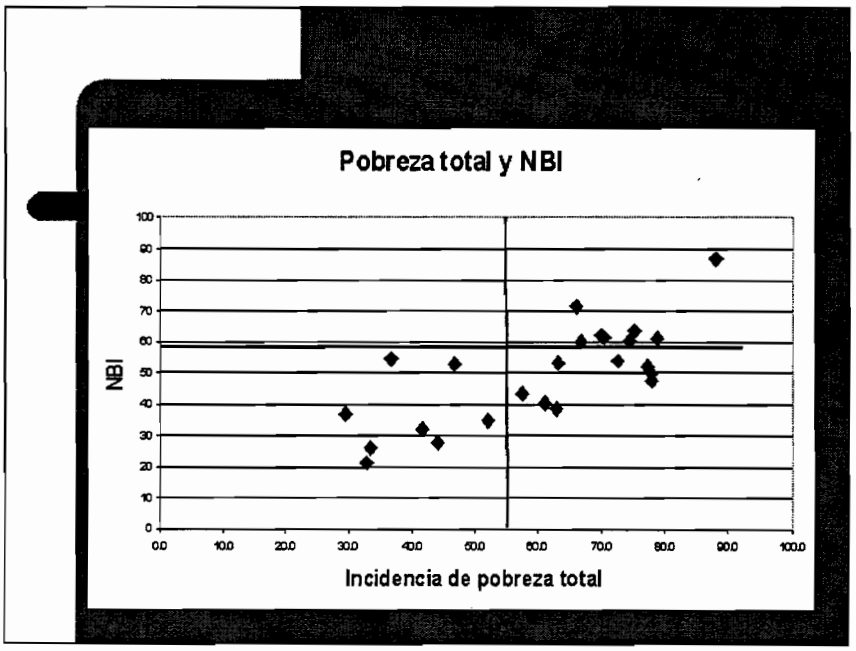
- Procedimientos estadísticos
  - Análisis exploratorio: identifica grupos afines muy rápidamente en función de una variable
  - Dalenius y Hodges: $(\sum fy)^{1/2})$
  - Diseños experimentales: cuadrados latinos, etc
  - Multidimensionales: Cluster


Indicadores de pobreza por departamentos 2001
Tasa de pobreza extrema (%)
Tasa de pobreza no extrema (%)
(%)
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Apurímac
Puno
Cajamarca
Cusco
Amazonas
Ayacucho
Huánuco
Loreto
Piura
Ancash
La Libertad
Tumbes
Arequipa
Ica
Madre de Dios
Lima y Callao
Tacna
Moquegua


Pobreza total: Incidencia y Severidad
Severidad
Incidencia


Pobreza total: Incidencia y población
Población pobre (%)
Incidencia (%)
20.0
15.0
10.0
5.0
0.0
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0


Pobreza extrema: Incidencia y población
Población pobre (%)
Incidencia (%)
15.0
10.0
5.0
0.0
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0



Diseños en etapas

- Los diseños óptimos: variables simples
- Precisión para una variable , imprecisión para otras.
- Los errores de muestreo en muestras de tamaños aceptables, son más pequeños que los errores no muestrales.

Métodos Para La Asignación De La Muestra

- Proporcional
- Optima de Neymann
- Potencial

Procedimientos Para La Selección De La Muestra

- Tipos de diseños.
  - Convencionales o clásicos: Selección previa al estudio
  - Adaptivos (para fenómenos raros o de pequeña prevalencia): Procedimientos de selección de la muestra en función de valores de la variable
- Tipos de selección.

Procedimientos Para La Selección de la muestra.

- Qué problemas están enfrentando los países en la estratificación de sus marcos?
  - Con el censo
  - Con los precensos.
  - Lecciones aprendidas

Muestras de panel

- Paneles para que?
  - Efectos de las pérdidas y rotaciones
  - Horizonte en tiempo del estudio de la dinámica de los procesos
  - Estudios retrospectivos y estudios prospectivos
- Viviendas?
- Hogares?
- Personas?
- Procedimientos para identificación de paneles:

Uso De Proyecciones De Población Para Generar Estimaciones

- Estimaciones puntuales de los totales poblacionales
  - "Cambios" en la tendencia por errores muestrales
  - Uso de "anclas", ajustes y balanceos

Procedimientos Para La Estratificación Del Marco De Muestreo

- Cuáles son las practicas que están teniendo los países en el diseño de sus muestras?
- Qué problemas están enfrentando los países en los diseños de sus muestras?
- Cuáles son las lecciones aprendidas para compartir.

SISTEMA DE INFORMACIÓN PARA LA SELECCIÓN DE LA MUESTRA

JUAN VALVERDE QUEZADA

INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA (INEI)
PERÚ

9° TALLER REGIONAL
Diseño y Construcción de Marcos de Muestreo
Para las Encuestas de Hogares

SISTEMA DE INFORMACION PARA LA SELECCIÓN DE LA MUESTRA

Lic. Juan Valverde Quezada
Instituto Nacional de Estadística e Informática - Perú

ESTRUCTURA DEL PRECENSO 1999

UNIDADES DE MUESTREO

UNIDADES PRIMARIAS DE MUESTREO (UPM)

- **UPM Urbana**

  Centro poblado con 2 mil y más habitantes.

- **UPM Rural**

  ** CCPP con 500 a menos de 2 mil habitantes,*
  ** Área de Empadronamiento Rural (AER)con menos de 500 habitantes (área dispersa).*

UNIDADES DE MUESTREO

UNIDADES SECUNDARIAS DE MUESTREO(USM)

- **USM Urbana**
  - *Sección Censal de una o más manzanas contiguas. Promedio: 80 viviendas*
- **USM Rurales**
  - *Sección Censal de una o más manzanas contiguas. Promedio: 80 viviendas*
  - *Viviendas particulares de las Áreas de Empadronamiento Rural (AER).*

UNIDADES DE MUESTREO

UNIDADES TERCIARIAS DE MUESTREO (UTM)

- **UTM Urbana**

  Viviendas particulares dentro de las USM urbanas seleccionadas.

- **UTM Rural**

  Viviendas particulares dentro de las USM rurales seleccionadas.

ESTRATIFICACIÓN DEL MARCO DE LA ENAHO

1ª Etapa: UPM

Área Urbana:

- **Estrato I:** UPM con 500 mil y más hab.
- **Estrato II:** UPM con 100 mil a 499 mil hab.
- **Estrato III:** UPM con 50 mil a 99 999mil hab.
- **Estrato IV:** UPM con 20 mil a 49 999 mil hab.
- **Estrato V:** UPM con 2 mil a 19 999 hab.

ESTRATIFICACIÓN DEL MARCO DE LA ENAHO

1ª Etapa: UPM

Área Rural:

- Estrato VI: UPM con 500 a 1 999 hab.
- Estrato VII: AER compuestos.
- Estrato VIII: AER simples

ESTRATIFICACIÓN DEL MARCO DE LA ENAHO

2ª Etapa: USM

- *USM nuevas en el marco (20%)*
- *USM antiguas (80%)*
- *USM en áreas periféricas*

SELECCIÓN DE LA MUESTRA

- En la ENAHO, la muestra es probabilistica, estratificada, multietapica e independiente en cada dominio de estudio.
- En la primera y segunda etapa se utiliza la selección sistemática con probabilidad proporcional al tamaño (PPT) de viviendas
- En la ultima etapa la selección es sistemática simple con arranque aleatorio.

SELECCIÓN DE LA MUESTRA

- **Lima Metropolitana**

1ra. Etapa	Autorepresentada
2ra. Etapa (Hoja de cálculo)	Selección Sistemática y proporcional al tamaño de viviendas.
3ra. Etapa (Computarizada)	Selección Sistemática Simple con arranque aleatorio

SELECCIÓN DE LA MUESTRA

- **Resto del País**

1ra. Etapa (Hoja de cálculo)

2ra. Etapa (Hoja de cálculo)

Selección Sistemática y proporcional al tamaño de viviendas.

3ra. Etapa (Computarizada)

Selección Sistemática Simple con arranque aleatorio.

CALCULO Y AJUSTE AUTOMATIZADO DE LOS FACTORES DE EXPANSION

Para que las estimaciones derivadas de la ENAHO, sean representativas de la población, es necesario multiplicar los datos de cada hogar muestral, por el peso o Factor de Expansión.

La probabilidad final de selección se calcula a través de la formula siguiente:

$$phij = \underbrace{\frac{(nh)(Mhi)}{Mh}}_{\textbf{1ra. etapa}} \; \underbrace{\frac{(ghi)(Mhij)}{Mhi}}_{\textbf{2da. etapa}} \; \underbrace{\frac{(mhij)}{M'hij}}_{\textbf{3ra. etapa}}$$

CALCULO Y AJUSTE AUTOMATIZADO DE LOS FACORRES DE EXPANSION

Primera Etapa de muestreo:

nh: Número de UPM a seleccionarse en el h-dominio.

Mh: Total de viviendas en el h-dominio.

Mhi: Total de viviendas en la i-UPM.

Segunda Etapa de Muestreo:

ghi: Número de USM a seleccionarse en la i-UPM del h- dominio.

CALCULO Y AJUSTE AUTOMATIZADO DE LOS FACTORES DE EXPANSION

Segunda Etapa de muestreo:

Mhij: Total de viviendas en la j-USM, de la i-UPM en el h-dominio.

Tercera Etapa de Muestreo:

mhij: Número de UTM a seleccionarse en la j-USM, de la i-UPM en el h- dominio.

M'hij: Total de viviendas listadas en la j-USM de la i-UPM en el h-dominio.

X. METODOLOGIA DE PONDERACION (Factor de Expansión)

- El Factor Básico de Expansión, es calculado como el inverso de la probabilidad final de selección (phij).

$$Whij = \frac{1}{phij}$$

X. METODOLOGIA DE PONDERACION (Factor de Expansión)

- Factor de Expansión Final (con Ajuste)

  Es importante ajustar los factores, teniendo en cuenta la magnitud de la no-respuesta.

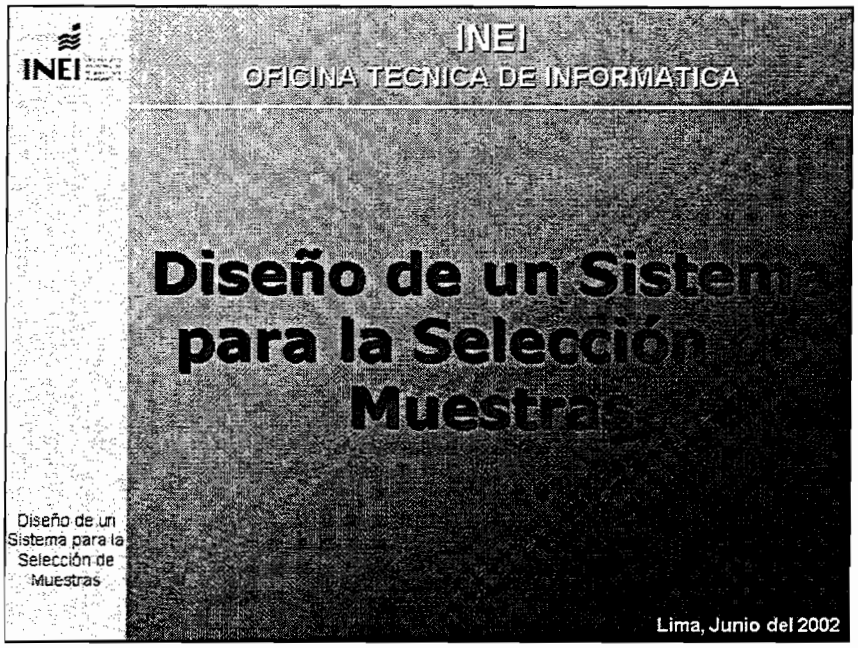
$$W'hij = (Whij) * \frac{(m'hij)}{(m''hij)}$$, donde:

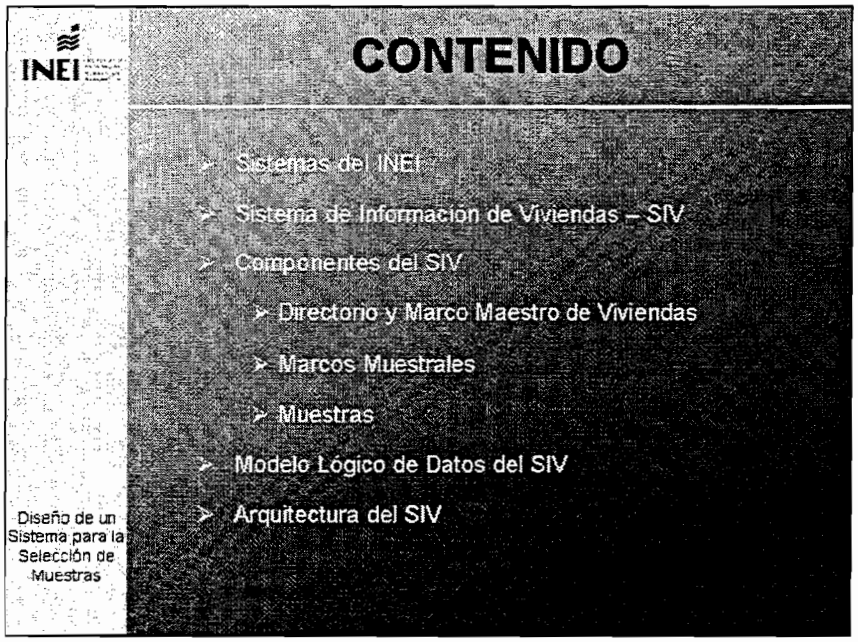
m'hij : viviendas seleccionadas

m''hij: viviendas encuestadas.

DISEÑO DE UN SISTEMA PARA LA SELECCIÓN DE MUESTRAS

INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA (INEI)
PERÚ


INEI
INEI
OFICINA TECNICA DE INFORMATICA
Diseño de un Sistema para la Selección de Muestras
Diseño de un Sistema para la Selección de Muestras
Lima, Junio del 2002


INEI
INEI
CONTENIDO
Sistemas del INEI
Sistema de Información de Viviendas – SIV
Componentes del SIV
Directorio y Marco Maestro de Viviendas
Marcos Muestrales
Muestras
Modelo Lógico de Datos del SIV
Arquitectura del SIV
Diseño de un Sistema para la Selección de Muestras

INEI

SISTEMAS INEI

- Sistemas Básicos
- Sistemas de Procesamiento de Censos y Encuestas
- Sistemas para la Difusión
- Sistemas para la Gestión

Diseño de un Sistema para la Selección de Muestras

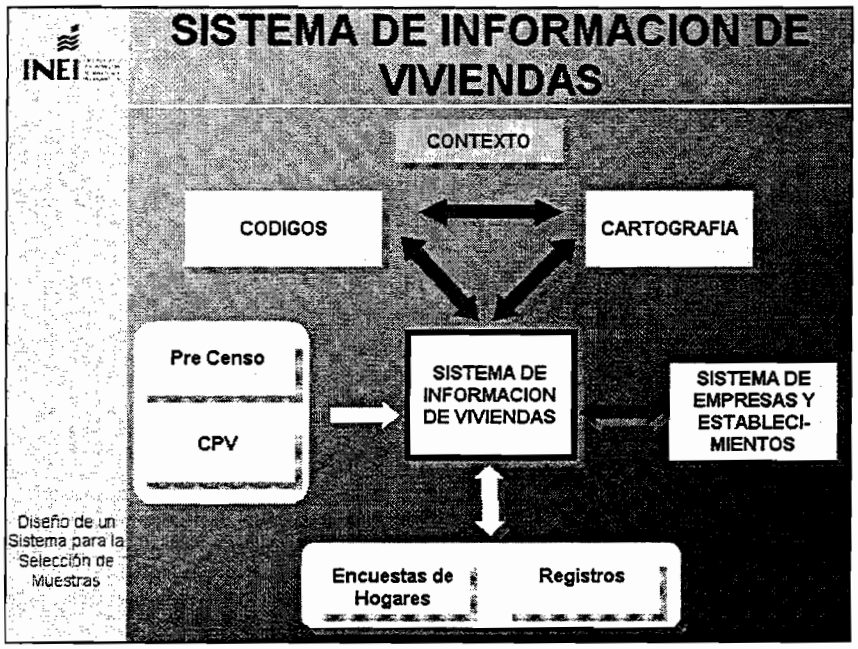
INEI

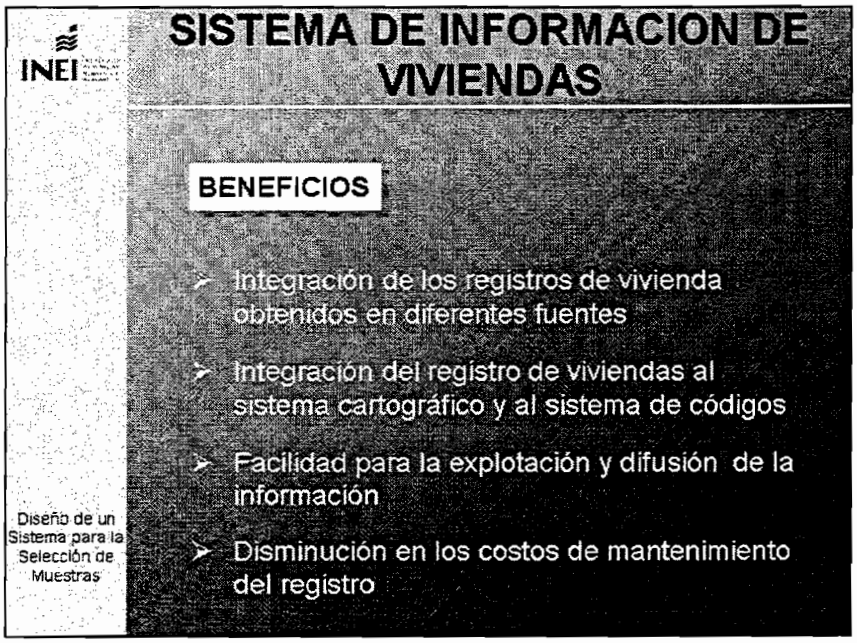
SISTEMA DE INFORMACION DE VIVIENDAS

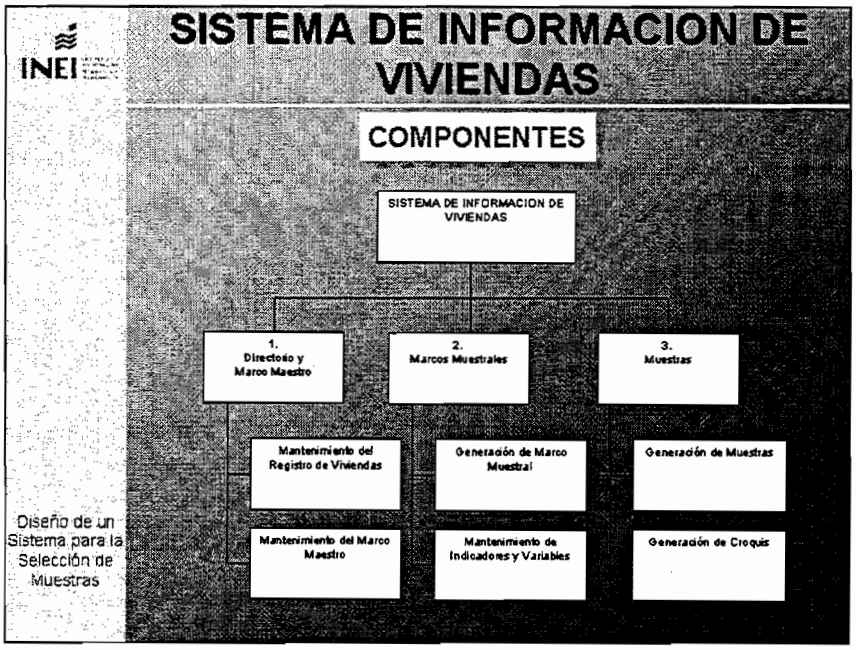
DEFINICION

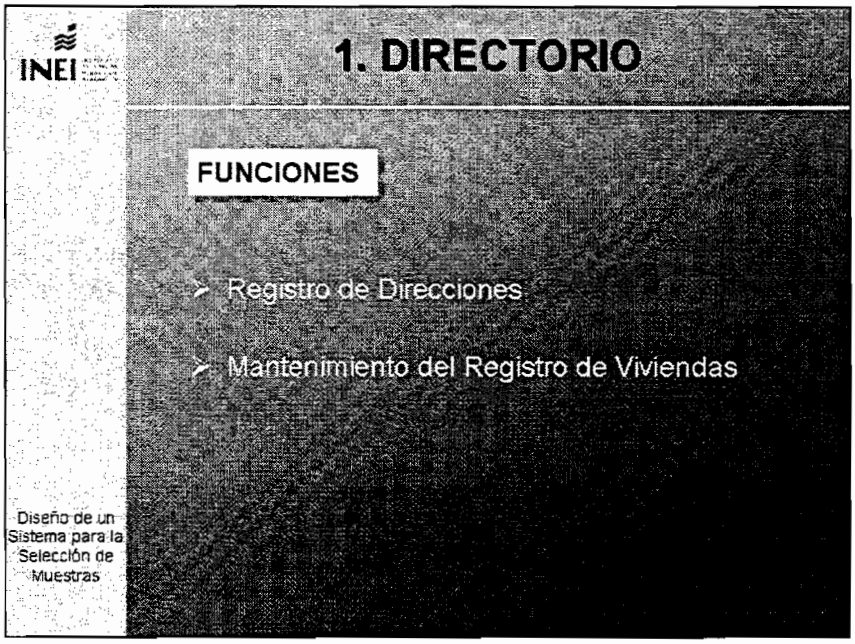
Es un sistema que registra la información de viviendas del país y permite su actualización y mantenimiento, así como la la generación del Marco Maestro y muestras de viviendas para la ejecución de encuestas.

Diseño de un Sistema para la Selección de Muestras


INEI
SISTEMA DE INFORMACION DE VIVIENDAS
CONTEXTO
CODIGOS
CARTOGRAFIA
Pre Censo
CPV
SISTEMA DE INFORMACION DE VIVIENDAS
SISTEMA DE EMPRESAS Y ESTABLECI-MIENTOS
Encuestas de Hogares
Registros
Diseño de un Sistema para la Selección de Muestras


INEI
SISTEMA DE INFORMACION DE VIVIENDAS
BENEFICIOS
Integración de los registros de vivienda obtenidos en diferentes fuentes
Integración del registro de viviendas al sistema cartográfico y al sistema de códigos
Facilidad para la explotación y difusión de la información
Disminución en los costos de mantenimiento del registro
Diseño de un Sistema para la Selección de Muestras


INEI
SISTEMA DE INFORMACION DE VIVIENDAS
COMPONENTES
SISTEMA DE INFORMACION DE VIVIENDAS
1. Directorio y Marco Maestro
2. Marcos Muestrales
3. Muestras
Mantenimiento del Registro de Viviendas
Generación de Marco Muestral
Generación de Muestras
Mantenimiento del Marco Maestro
Mantenimiento de Indicadores y Variables
Generación de Croquis
Diseño de un Sistema para la Selección de Muestras


INEI
1. DIRECTORIO
FUNCIONES
Registro de Direcciones
Mantenimiento del Registro de Viviendas
Diseño de un Sistema para la Selección de Muestras

INEI

1. MARCO MAESTRO

FUNCIONES

- Construcción del Marco Muestral
- Estratificación del Marco Muestral
- Administración de Tablas Auxiliares
- Generación de Reportes

Diseño de un Sistema para la Selección de Muestras

INEI

2. MARCOS MUESTRALES

FUNCIONES

- Construcción de distintos Marcos Muestrales
- Definición de diferentes Etapas de Muestreo para cada Marco Muestral
- Cada Etapa de Muestreo se define a partir de las entidades geográficas administrativas
- La Unidad Muestral puede ser una entidad individual o una agrupación de éstas
- Numeración de Unidades Muestrales de acuerdo a criterios definidos por el usuario
- Se utiliza la definición muestral de área urbana y rural para la definición de la Unidad Muestral

Diseño de un Sistema para la Selección de Muestras

INEI

2. MARCOS MUESTRALES

FUNCIONES

- Estratificación del Marco Muestral a partir de variables asociadas a la entidad geográfica administrativa, base de la Unidad Muestral
- Estratificación en las diferentes Etapas de Muestreo
- Diferentes Estratos pueden contener a la misma Unidad Muestral
- Definición de Estratos sugerido por el Sistema, con opción de reestructuración por parte del usuario
- Reestructuración, a criterio del usuario, de Estratos ya definidos en el sistema por razones de eficiencia

Diseño de un Sistema para la Selección de Muestras

INEI

2. MARCOS MUESTRALES

FUNCIONES

- Administración de información auxiliar
- Generación de Reportes
  - Directorios de Unidades Muestrales
  - Resúmenes
  - Reportes de Control

Diseño de un Sistema para la Selección de Muestras

INEI

3. MUESTRAS

FUNCIONES

- Cálculo del tamaño de la muestra
  - Utilización de estimaciones anteriores de las variables de estudio
  - Selección del nivel de inferencia y de la fórmula correcta a utilizar
- Selección de la muestra
  - Integración con las Unidades Muestrales del Marco
  - Selección de la muestra en las área urbana y rural simultáneamente
  - Diferentes métodos de selección (modificables)
  - Reporte de verificación de selección

Diseño de un Sistema para la Selección de Muestras

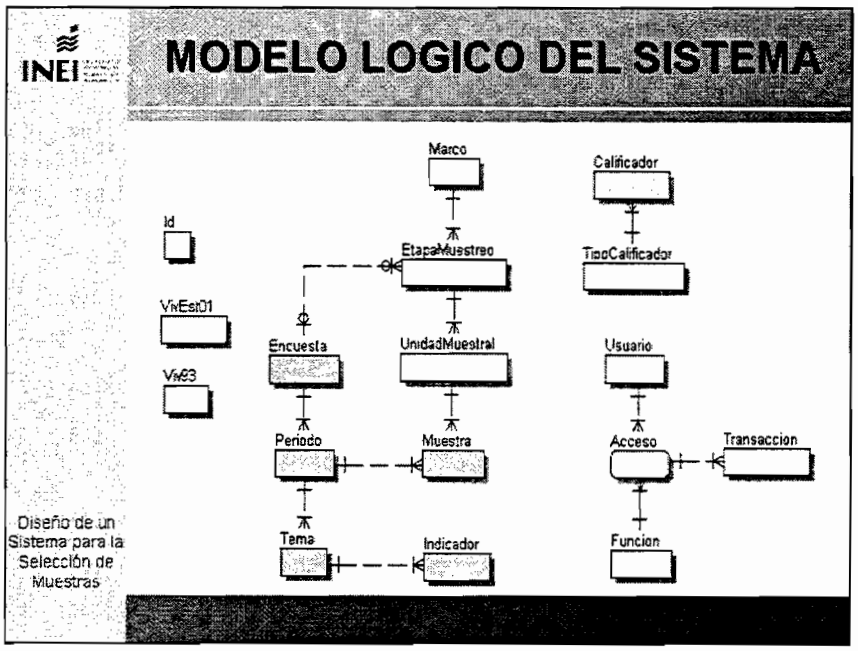
INEI

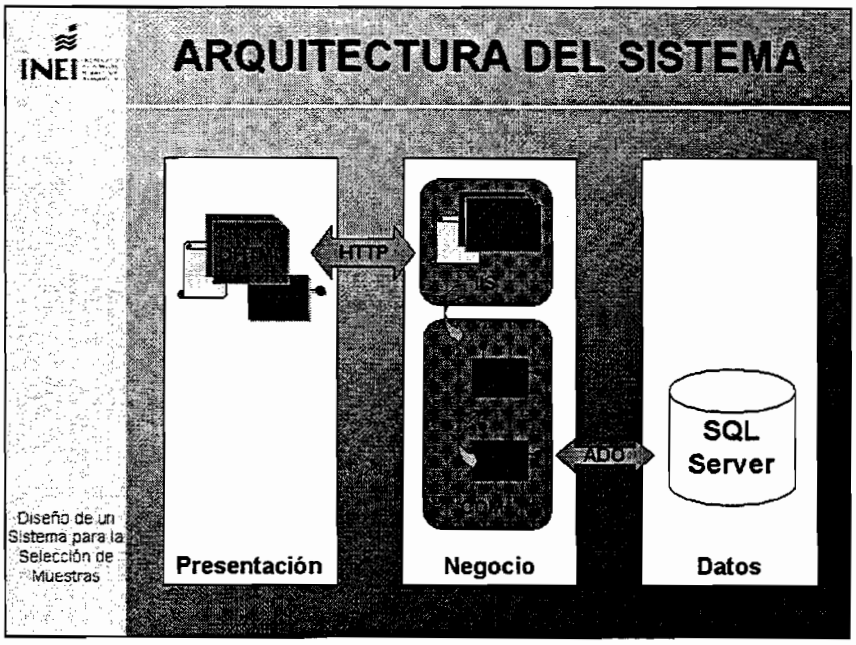
3. MUESTRAS

FUNCIONES

- Muestra maestra y muestra panel
- Generación de croquis automatizado
- Administración de encuestas y muestras
- Administración de indicadores
- Reportes de unidades muestrales seleccionadas

Diseño de un Sistema para la Selección de Muestras


INEI
MODELO LOGICO DEL SISTEMA
Marco
Calificador
Id
EtapaMuestreo
TipoCalificador
VivEst01
Encuesta
UnidadMuestral
Usuario
Vv63
Periodo
Muestra
Acceso
Transaccion
Tema
Indicador
Funcion
Diseño de un Sistema para la Selección de Muestras


INEI
ARQUITECTURA DEL SISTEMA
HTTP
IIS
ADO
SQL Server
Presentación
Negocio
Datos
Diseño de un Sistema para la Selección de Muestras

MARCOS DE ÁREAS PARA ENCUESTAS COMPLEJAS DE VIVIENDAS

ALFREDO ALIAGA
CEPAL

ÍNDICE

Página

1.- Importancia de los Marcos de Áreas para la Selección de Viviendas 95
2.- Requerimientos de Información, Dominios de Estudio e Inferencia Estadística 95
3.- Marcos Maestros y Diseño de Muestras Maestras 96
4.- Procedimientos para la Formación de las Unidades del Marco de Muestreo 96
5.- Tamaño de Unidades Primarias de Muestreo y su Efecto
en la Varianza de los Estimadores 96

1. Importancia de los Marcos de Áreas para la Selección de Viviendas

Objetivos

- Estudiar diferentes unidades de población
- A través de la población localizada en viviendas (hogares) seleccionadas
- A través de viviendas localizadas en un tipo de muestreo de conglomerados (áreas)

Características de una Lista de Viviendas

- o No existe una lista de personas y/o de viviendas completa y/o recién actualizada
- o Si tal lista existiese quizás no seria útil por su inestabilidad de permanencia

- Características de una Lista de Áreas
  - o Áreas incluyen y cubren todas las viviendas
  - o Cada área tiene material cartográfico (bosquejo o mapa)
  - o Cada área tiene limites permanentes entre el momento de la actualización y del trabajo de campo
- Aspectos importantes
  - o Se requiere:
    - o Lista de áreas
    - o Cartografía de cada área
    - o Medida de cada área
  - o El aspecto cartográfico de cada área se convierte en el factor primario de este tipo de muestreo
  - o En muchos países una lista de áreas (segmentos) censales constituye una lista de áreas

2. Requerimientos de Información, Dominios de Estudio e Inferencia Estadística

- Las técnicas descritas en los cursos de estadística asumen que las observaciones son independientes e idénticamente distribuidas y en general estos supuestos están definidas en poblaciones infinitas.
- Información que provienen de encuestas no cumplen con dichas condiciones ya que provienen de poblaciones finitas y con características (variables) categorizadas.
- Entonces cuales son los efectos de dichas técnicas en poblaciones finitas

Efectos de comparabilidad entre dominios como comparar dos dominios diferentes

Efectos en los intervalos de confianza

Ejemplo:

en poblaciones infinitas

$$p + (-)\ 2 \times es_{srs}$$

en poblaciones finitas

$$p + (-)\ 2 \times deff \times es_{srs}$$

Efectos en las pruebas del chi-cuadrado crece el valor del nivel de significación

Efectos en el análisis de regresión Uso de las ponderaciones o no

3. Marcos Maestros y Diseño de Muestras Maestras

Requerimientos de un marco maestro para un muestreo bi-etapico

Lista de áreas de tamaño 300-500 hhs

Cartografía de cada área (censal)

En general creado en base del ultimo censo y renovable cada censo

Requerimientos de una muestra maestra

Es una selección de una muestra grande de áreas

Debe ser bastante flexible para la consideración de dominios en las encuestas futuras basadas en ellas

Actualización programada de acuerdo a como se este usando la muestra maestra

4. Procedimiento para la Formación de las Unidades del Marco de Muestreo

Asumiremos que el muestreo es bietapico
Definición exacta de los limites del área total que cubre la población a ser encuestada

Áreas a ser omitidas

Por ejemplo la población que residen en hogares residenciales en la ciudad X
En muchos países, las áreas censales de supervisión sirven como base para la elaboración de las unidades muestrales

- Los limites deben ser claros y exactos
- Estratificación de las unidades muestrales
- Estratificación geográfica
- Con la numeración de las unidades, por ejemplo: en forma serpentina

5. Tamaño de Unidades Primarias de Muestreo y su Efecto en la Varianza de los Estimadores

Depende del numero de etapas de muestreo

Menor el numero de etapas mejor control del error muestral

Por ejemplo en muestreo tri-etapico:

Error muestral = $SS_1/a + SS_2/ab + SS_3/abc$

Tamaño depende del factor costo, el conglomeramiento de las unidades en sus diferentes etapas reduce el costo pero incrementa el error muestral

$$\text{Error muestral} = \text{deff x es}$$

En general el efecto de las unidades primarias en el error muestral es que provee la mayor componente en el valor error muestral.

Cuando el muestreo es tri-etapico:

Por ejemplo: se tiene un país con 3000 distritos

Se estratifica distritos vecinos en estratos de tamaño de 500,000 habitantes. Total poblacional es 250 millones de habitantes

Se seleccionan un solo distrito en cada estrato, en total se selecciona 500 distritos

En cada distrito seleccionado se continua con un muestreo bi-etapico

Optimalidad se define claramente en un muestreo bi-etapico, en un muestreo tri-etapico el tamaño de las unidades primarias pueden ser de muy diferentes tamaños controlándose con selección PPT

Tamaño optimo depende:

de una función de costo

$$C = C_0 \sqrt{m} + C_1 m + C_2 m n$$

minimizar la varianza con

$$DEFF = 1 + \delta (n - 1)$$

tamaño optimo: opt n =

$$\sqrt{[(C_1 + C_0/2 \sqrt{m})/C_2)]x[(1-\delta)/\delta]}$$

INFORME TÉCNICO: EL MARCO DE MUESTREO NACIONAL DE VIVIENDAS Y LA ENCUESTA PERMANENTE DE HOGARES: ESTADO ACTUAL Y SU REFORMULACIÓN

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS (INDEC)
ARGENTINA

ÍNDICE

Página

1.- El Marco de Muestreo Nacional de Vivienda de Argentina 102
1.1 Características del marco actualmente disponible 102
1.2 Necesidades de extender su cobertura 102
1.3 Objetivos de la extensión 103
1.4 Lineamientos generales del diseño de muestra para la extensión del actual marco 103
1.4.1 Primer etapa de muestreo 104
1.4.2 Segunta etapa de muestreo 104
1.5 Implementación del nuevo marco de muestreo 105
1.5.1 Implementación en las áreas rurales 106
1.5.2 Implementación en las áreas urbanas 106
1.6 Sistema Informático de soporte del marco 107
2.- La Encuesta Permanente de Hogares de Argentina 107
2.1 Características generales del diseño muestral de la EPH 107
2.1.1 Ámbito geográfico 107
2.1.2 Periodicidad 108
2.1.3 Marco muestral 108
2.1.4 Tipo de muestreo 108
2.1.5 Estratificación 108
2.1.6 Tamaño de la muestra 108
2.1.7 Rotación de la muestra 109
2.1.8 actualización de los listados de viviendas 109
2.1.9 Estimaciones 109
2.2 Reformulación muestral de la EPH 110
2.2.1 Captación de cambios intraanuales 110
2.2.2 Ampliación de la cobertura geográfica 110
2.3 Principales características de la EPH Continua 110
2.3.1 Dominios de estimación 110
2.3.2 Tamaño de la muestra 111
2.3.3 Asignación de la muestra 111
2.3.4 Diseño de la muestra 112
2.3.5 Asignación de las UPEs a las semanas del trimestre 112

2.3.6 Esquema de rotación 112
Anexo A 114
Anexo B 115

1 El Marco de Muestreo Nacional de Viviendas de Argentina

El plan de acción del programa MECOVI - ARGENTINA contempla, en la definición de la estrategia a mediano plazo para el diseño de un Sistema Integrado de Encuestas a Hogares (SIEH), lograr como meta una cobertura nacional urbana y rural para las encuestas a desarrollar por el INDEC, dentro de dicho sistema. Para ello se está trabajando en la ampliación hacia las áreas rurales del marco muestral actualmente utilizado para las encuestas a hogares.

En el presente documento se describen brevemente las características del marco actual, se analizan las necesidades de dicha ampliación, sus objetivos y los procedimientos que se están llevando a cabo para lograr estos objetivos.

1.1 Características del marco actualmente disponible

El Marco de Muestreo Nacional Urbano (MMNU), actualmente disponible en el INDEC, responde al concepto de marco de muestreo principal; se trata del listado completo de direcciones de viviendas, con su correspondiente ubicación cartográfica, de una muestra de áreas urbanas del país. A partir de este listado se obtienen muestras de viviendas para diversas encuestas.

El MMNU cubre la población urbana residente en localidades que, según el Censo Nacional de Población y Vivienda de 1991, tenían 5000 habitantes o más. Esta población representaba, en 1991, el 83 % de la población total del país.

La muestra de áreas que componen el Marco, se ha obtenido en dos etapas de muestreo:

1) en la primer etapa se han muestreado, probabilísticamente, unidades primarias definidas como aglomeraciones (centros urbanos y localidades unidas por continuidad urbana),
2) en la segunda etapa se han muestreado, dentro de cada unidad primaria seleccionada y también probabilísticamente, unidades secundarias o áreas definidas a base de radios censales (conjunto de manzanas con alrededor de 300 viviendas).

Este procedimiento en dos etapas permite un ahorro notorio de recursos, restringiendo a un limitado espacio territorial el trabajo de listar viviendas y mantener actualizados estos listados. Permite, a la vez, el ahorro de recursos en cada encuesta al concentrar las tareas de recolección de datos y posibilita una adecuada organización operativa para las actividades de supervisión.

Para las aglomeraciones donde se realiza la Encuesta Permanente de Hogares (son 28 aglomeraciones autorrepresentadas), se incorporaron al MMNU la totalidad de áreas que componían sus muestras. Estas áreas fueron consideradas como unidades secundarias seleccionadas.

La población no cubierta por el Marco actual, se distribuye en pequeñas localidades y en zonas rurales. En algunas provincias (Mendoza, San Juan, Misiones) se concentra en zonas semiurbanas lindantes a localidades urbanas.

1.2 Necesidades de extender su cobertura.

Las necesidades de una extensión están dadas , por una parte, por la falta de cobertura total a nivel nacional para encuestas cuya población objetivo sea la población total del país; la falta de cobertura total genera, también, dificultades en la comparación de resultados de encuestas análogas entre países, en particular entre países del MERCOSUR. Por otra parte, en el caso de encuestas a hogares con dominios de análisis provinciales, es necesaria una extensión a las zonas rurales para mejorar la representatividad de las provincias con poca población.

Es importante tener presente que, al definir como dominio de análisis a cada provincia, habrá que sobremuestrear aquellas que tienen muy bajas densidades de población, con relación a las más pobladas.

Este desbalance se acentúa en razón de que las zonas más despobladas son relativamente más costosas de relevar en campo.

Por lo tanto, se debe considerar caso por caso la solución más adecuada, incluyendo la delimitación operativa del universo a muestrear, sobre la base de criterios lo más uniformes posibles para preservar, al mismo tiempo, el nivel nacional y la comparabilidad entre provincias.

1.3 Objetivos de la extensión

Los objetivos específicos de la extensión del MMNU son:

1. posibilitar la extracción de muestras para encuestas a hogares con cobertura nacional, incluyendo tanto a población urbana como rural,
2. posibilitar la extracción de muestras para diversos dominios de análisis: nacional, regional y provincial; la asignación de los tamaños de muestra de áreas resultantes estarán limitados por los costos esperados de relevamiento y por el balance entre los requerimientos de muestra a nivel nacional y provinciales,
3. brindar flexibilidad en el uso del Marco, de modo de facilitar los mecanismos de obtención de muestras para dominios de análisis con mayor desglose geográfico (provincias) y de submuestras de áreas para dominios más agregados (total del país y regiones).

1.4 Lineamientos generales del diseño de muestra para la extensión del actual marco

Del análisis de las alternativas planteadas para la extensión del actual marco urbano, surgió la conveniencia de ampliarlo redefiniendo la totalidad de unidades primarias, en vez de muestrear de un modo complementario las zonas no incluidas en el MMNU.

Se han delineado las características y las etapas de elaboración de un nuevo marco de muestreo para encuestas a hogares con cobertura nacional total que, a la vez, utilice el máximo posible de áreas ya incluidas en el actual MMNU, de modo de ahorrar recursos en la nueva implementación.

Tal como el anterior MMNU, el nuevo marco contendrá la identificación completa de las viviendas de una muestra probabilística de áreas. Este muestreo de áreas se realiza en dos etapas: las unidades primarias (UPM) están configuradas, en general, por departamentos o conjunto de ellos (partidos en el caso de la provincia de Buenos Aires); las unidades secundarias (USM) se configuran con radios censales. En algunos casos de excepción, las UPM se han definido tomando como base las fracciones censales.

El universo de referencia de este nuevo marco es la totalidad de viviendas del país y cada provincia es considerada como un dominio de análisis independiente.

Del territorio total del país se han eliminado las zonas inaccesibles y aquéllas de alto costo de relevamiento por su baja densidad de población. Las zonas eliminadas corresponden a extensas superficies despobladas, alejadas de los centros urbanos; en general, no incluyen pequeños centros poblados o caseríos y tienen por radio censal, según los datos del Censo de 1991, una densidad de viviendas particulares ocupadas menor a una cada 1.000 ha., con menos de 50 viviendas totales.

Con el propósito de fijar la cantidad total de áreas a incluir en el marco se ha tenido en cuenta un esquema básico de tamaños de muestra compatible con los requerimientos de las encuestas más usuales que viene realizando el INDEC. Este esquema se basa en muestras nacionales del orden de 36.000 viviendas, con las siguientes asignaciones:

- una asignación equiprobable del total nacional por provincia, salvo en las provincias con menor cantidad de población para las cuales se ha asignado un mínimo de 1000; en estos casos el sobremuestreo se compensa con una asignación menor a la equiprobable en el Gran Buenos Aires.

- la cantidad de unidades secundarias a seleccionar, por cada unidad primaria no autorrepresentada, es de 12.
- la cantidad de viviendas a seleccionar por unidad secundaria, para una muestra nacional de 36.000 viviendas, se calcula que serán 10 (eventualmente 5 en las localidades urbanas de mayor tamaño).
- se ha tratado, en lo posible, que la cantidad de UPM a seleccionar como no autorrepresentadas sea par en cada dominio de análisis.

La información y la cartografía básicas utilizadas para la elaboración de la muestra de áreas es la del Censo Nacional de Población y Viviendas de 1991. Para la segunda etapa se han solicitado a las Direcciones Provinciales de Estadística algunos datos complementarios de las actualizaciones realizadas para el Censo de 2001.

Como ya se mencionó, las muestras de áreas se elaboran independientemente para cada provincia.

1.4.1 Primer etapa de muestreo

La estratificación de las UPM se basa en :

- el tamaño de la localidad con mayor población de la UPM,
- la proporción de población rural,
- la ubicación geográfica dentro de la provincia y la distancia entre la localidad más importante y la capital de la provincia,

Los dos primeros criterios se han utilizado para la conformación de estratos explícitos y el tercero para elaborar patterns de selección controlada. En cada caso se han definido tantos estratos explícitos como cantidad de UPM a seleccionar.

La probabilidad asignada a cada UPM es proporcional a una medida de tamaño establecida como la cantidad de viviendas particulares ocupadas, según el Censo de 1991.

A cada UPM se le ha asignado una probabilidad condicional, teniendo en cuenta la pertenencia o no al MMNU, de modo que:

- la probabilidad total sea proporcional a su medida de tamaño;
- sea máxima la cantidad de UPM seleccionadas que contengan áreas del MMNU, es decir, que sea máxima la superposición entre el nuevo marco y el MMNU.

Para el cálculo de estas probabilidades condicionales se revisaron y ensayaron varios procedimientos. Finalmente se aplicó un procedimiento similar al descripto en "CPS Technical Paper 63", Appendix A, 6/4/97, y que tiene en cuenta el diseño originario del MMNU.

1.4.2 Segunda etapa de muestreo

Dentro de cada UPM seleccionada se han revisado los radios censales que no habiendo sido excluidos en la primer etapa, presentan problemas operativos, ya sea por la baja densidad de viviendas, o por su orografía y vías de comunicación.

Cada USM se ha configurado con uno o, en su defecto, dos o más radios censales contiguos, de modo que contenga al menos alrededor de 50 viviendas. Se ha controlado que su superficie total no exceda 45.000 has. Cuando la superficie de los radios censales que configuran un área es muy extensa, ésta se ha delimitado excluyendo las zonas más inaccesibles y en donde se espera una menor cantidad de viviendas.

Las USM se estratifican a base de indicadores similares a los utilizados en la elaboración del MMNU, eliminando aquéllos que tuvieron un peso relativo escaso en las clasificaciones de dicha elaboración. Estos indicadores se reducen a:

- porcentajes de viviendas con agua por cañería de fuente pública o bomba a motor;
- porcentaje de viviendas sin piso de tierra;
- porcentaje de hogares con baño con descarga de agua de uso exclusivo;
- porcentaje de jóvenes de 15 a 29 años con primaria completa;
- porcentaje de jóvenes de 20 a 34 años con secundaria completa

A estos se agregan los siguientes:

- tamaño de la localidad de pertenencia total o parcial,
- distancia estimada media de la USM a la localidad más importante de la UPM,

El proceso de estratificación se realiza automáticamente y del siguiente modo:

El primer criterio de estratificación está dado por los indicadores de características de población y vivienda según el censo de 1991, y obtenido a través de técnicas multivariadas. La cantidad de clases a definir es exactamente la cuarta parte de la cantidad de USM a seleccionar, hasta un máximo de 12, y de modo que cada clase contenga la misma cantidad de USS (USM seleccionadas); es decir cada estrato acumula la misma cantidad de medidas de probabilidad asignadas.

Este modo de fijar las cantidades de estratos se debe a la necesidad de controlar, de un modo ágil, la asignación balanceada de submuestras por trimestre y por mes para aquellas encuestas que se desarrollen a lo largo de un año, o por semana para las que se desarrollan a lo largo de un mes. Por ejemplo, en el caso de las UPS con probabilidades menores a 1, de cada estrato por nivel socioeconómico se seleccionan 4 USM.

Para el segundo criterio de estratificación, las USM se ordenan de acuerdo al tamaño de la localidad principal contenida en ellas o a la que pertenecen total o parcialmente, seguidas de las USM rurales ordenadas por su número de identificación, cuya secuencia corresponde a la numeración consecutiva de fracciones y radios. Siguiendo este orden se definen tantos estratos como USM a seleccionar.

Por último, se utiliza como factor de control en la selección la distancia de cada USM a la localidad principal de la UPS o alguna otra característica geográfica.

En el caso de las UPS con probabilidades menores a 1 se seleccionan 12 USM. En el caso de las UPS autorrepresentadas las cantidades de USS y los estratos, correspondientes al primer criterio de clasificación, dependen de la cantidad de viviendas que les corresponde en una muestra estándar.

La medida de tamaño asignada a cada USM es la cantidad de viviendas particulares ocupadas.

La parte proporcional de USM a seleccionar en localidades incluidas en el MMNU, se obtiene submuestreando áreas del MMNU teniendo en cuenta el estrato de pertenencia.

1.5 Implementación del nuevo marco de muestreo

Las últimas etapas de elaboración del nuevo Marco ampliado consisten en identificar en terreno, y luego registrar en un sistema informático, todas las viviendas existentes en cada una de las nuevas áreas incorporadas. La primer tarea se realiza actualizando la cartografía de cada área y listando exhaustivamente sus viviendas. Finalmente estos datos son incorporados a un sistema computacional que permite realizar la extracción de muestras para las diversas encuestas, así como emitir los archivos necesarios y obtener los materiales para el trabajo en campo: hojas de ruta, etiquetas de formularios, etc.

Las USM se clasifican en cuatro categorías a tener en cuenta para el trabajo en campo de implementación del marco y de las futuras encuestas. Estas categorías, que definen modalidades operativas distintas, son las siguientes:

1. USM incluidas totalmente en una localidad.
2. USM con población dispersa y con población en una localidad que incluye totalmente al menos otra USM.
3. USM con población dispersa y que contienen una o más localidades, completas o incompletas, pero sin otras USM totalmente contenidas en ellas.
4. USM que contienen sólo población dispersa.

En los casos 2 y 3, para asignar la modalidad operativa de trabajo de campo, también se tiene en cuenta el porcentaje de viviendas particulares ocupadas en zonas amanzanadas o nucleadas (urbanas o en pueblos o caseríos),

En general, en el caso 2 el trabajo de campo se realiza como habitualmente se hace en zonas urbanas (categoría 1) y la selección de muestras del marco para las futuras encuestas se hará por viviendas. En el caso 3, el trabajo de campo corresponde al de zonas rurales (categoría 4) y la selección se hará por segmentos de viviendas.

1.5.1 Implementación en las áreas rurales

Para la realización del trabajo de campo en las zonas rurales, el INDEC firmó un convenio con el Instituto Nacional de Tecnología Agropecuaria (INTA), ya que el INTA es una institución con una red de servicios que cubre todo el país y, por otra parte, porque el INDEC mantiene un convenio de colaboración técnica con dicha entidad.

La implementación en campo de las áreas rurales comenzó en julio de 2001 y se está llevando a cabo por regiones, en forma escalonada. En ese mismo año se realizó la actualización cartográfica para el Censo de Población 2001 pero sólo en localidades. Por lo tanto no se dispone de cartografía rural actualizada para realizar el listado de viviendas. Por ello, entre las tareas del trabajo de campo, se contempla la actualización cartográfica de las áreas rurales seleccionadas.

Como herramienta de trabajo en las áreas rurales se ha incorporado el uso de GPS (Geographical Positioning System) de manera de posicionar las coordenadas geográficas de viviendas y puntos de referencias, útiles para la ubicación de las viviendas que resulten seleccionadas en futuras encuestas a hogares.

Con el archivo de datos obtenidos a través del GPS será posible dibujar esquemas de las áreas y ubicar estos puntos en dichos esquemas.

Hasta la fecha se han finalizado las tareas de campo en las zonas rurales de la región Cuyo y de las provincias de Salta, Jujuy y Tucumán. En estos días se está llevando a cabo el trabajo de campo en las provincias de Santiago del Estero, La Rioja y Catamarca.

1.5.2 Implementación en las áreas urbanas

Para obtener los listados de viviendas en las nuevas áreas urbanas, las tareas de campo serán llevadas a cabo por personal de las Direcciones Provinciales de Estadística, tal cual lo han hecho en las áreas del MMNU y en las aglomeraciones de la Encuesta Permanente de Hogares.

En este caso, sí se dispone de cartografía digitalizada actualizada en el año 2001, tarea previa a la realización del Censo de Población de ese año.

1.6 Sistema informático de soporte del marco

Para el nuevo marco de viviendas, al igual que se hizo para el marco urbano, se está diseñando un sistema informático de soporte que cumpla con los siguientes objetivos:

Objetivos atinentes a la obtención de muestras de viviendas:

- posibilitar la selección de viviendas para cada encuesta en particular, mediante la elección de rutinas de procedimientos alternativos;
- generar archivos de datos y emitir listados para las tareas de campo de cada encuesta;
- registrar los usos del marco: encuestas, fechas, fracciones de selección, viviendas seleccionadas en cada caso;
- generar archivos con los factores de ponderación de cada área para encuestas con muestras no autoponderadas, de modo de facilitar los procesos de estimación de resultados.

Objetivos atinentes a la implementación y mantenimiento del marco:

- disponer de un sistema de ingreso, consistencia de datos y modificaciones;
- facilitar la emisión de listados y archivos de impresión para actualizaciones periódicas;
- registrar las fechas de actualización;
- emitir salidas que faciliten el control de los procesos realizados;
- emitir resúmenes e indicadores sobre características de las unidades del marco;
- posibilitar la introducción de cambios necesarios a lo largo del tiempo.

En función de estos objetivos, se está desarrollando un sistema global con varios subsistemas que intercambian información entre sí, con soporte lógico sobre INFORMIX y SAS.

2 La Encuesta Permanente de Hogares de Argentina

El principal usuario del actual Marco de Muestreo Nacional Urbano es un programa de carácter permanente, la denominada Encuesta Permanente de Hogares (EPH)

La EPH surgió en la década del 70, como un intento de dar respuesta a las necesidades de información sobre distintas problemáticas sociales, principalmente las relacionadas con el mercado de trabajo, en los principales centros urbanos del País.

2.1 Características generales del diseño muestral actual de la EPH

2.1.1 Ámbito geográfico

La EPH se realiza en:

- El Gran Buenos Aires
- En todos los aglomerados urbanos de más de 100.000 habitantes o menores pero que son capitales de provincias (con la excepción de San Nicolás)
- El aglomerado urbano - rural Alto Valle del Río Negro.

Los dominios de estimación son:

- Cada uno de los 28 aglomerados
- Conjunto de los 28 aglomerados

- El aglomerado urbano - rural del Alto Valle del Río Negro

2.1.2 Periodicidad

Dos mediciones anuales, en general en los meses de Mayo y Octubre

2.1.3 Marco muestral

La República Argentina esta dividida desde el punto de vista político-administrativo en 23 provincias y un distrito federal (Ciudad Autónoma de Buenos Aires)

Cada provincia está dividida en unidades políticas menores : departamentos o partidos. Estos, a los fines censales, se subdividen a su vez en: fracciones, y estas en radios censales. Cada radio censal está constituido por un conjunto de aproximadamente 300 viviendas.

El marco muestral de cada localidad que integra la encuesta se elaboró a partir de todos los radios definidos por el Censo Nacional de Población y Vivienda para dicha localidad

La EPH trabajó con la información censal a nivel de radio *y a nivel de manzana dentro del radio* (cuando fue necesario subdividirlos debido a su tamaño en viviendas, o a su extensión). Por primera vez, a partir del Censo Nacional de 1991, los muestristas tuvieron un acceso directo a las bases censales. Se contó con información sobre viviendas particulares habitadas, tipo de vivienda, servicios de la vivienda, población por sexo, cantidad de jefes de hogar y sus características principales, información provista por el cuestionario Básico.

2.1.4 Tipo de muestreo

En general, el diseño muestral de la EPH consiste en muestras bietápicas estratificadas, es decir que las unidades primarias (radios censales o subdivisiones de los mismos) son agrupadas en estratos, para luego realizar una selección de las mismas con probabilidad proporcional a su tamaño, medido en cantidad de viviendas particulares.

En una segunda etapa se seleccionaron las viviendas en forma sistemática, dentro de las unidades primarias seleccionadas, a partir de un listado de viviendas elaborado en cada una de esas unidades. En el caso del Alto Valle de Río Negro se utilizó una muestra trietápica en la zona rural y una muestra bietápica estratificada en la zona urbana.

2.1.5 Estratificación

Las unidades de primera etapa fueron estratificadas según las siguientes variables provistas por el Censo de Población 1991:

- Porcentaje por unidad primaria de jefes que nunca asistieron a establecimiento educativo más el porcentaje de jefes que tienen nivel educativo "Primario incompleto" (Y)
- Porcentaje por unidad primaria de jefes con nivel educativo "Universitario completo" (X)

2.1.6 Tamaño de la muestra

A los efectos del diseño los aglomerados fueron ordenados en 5 grupos de acuerdo a la cantidad de viviendas que tenían al momento del Censo 91. En el siguiente cuadro se muestra la cantidad de unidades de primera etapa (UPE) a seleccionar dentro de cada grupo, así como la cantidad de unidades secundarias (USE) a seleccionar dentro de cada UPE. (para el MMNU las UPEs pasaron a ser unidades de segunda etapa cuando estos diseños se integraron al mismo en el año 1995. Cada aglomerado EPH fue considerado como una unidad primaria autorrepresentada)

Grupo	Tamaño del aglomerado en viviendas censo 1991	UPEs a seleccionar	USEs por UPE	Tamaño de la muestra
1	X ≥ 3.000.000	524	8	4192
2	3.000.000 > X > 300.000	100	10	1000
3	80.000 ≤ X< 300.000	80	10	800
4	40.000 ≤ X< 80.000	60	12	720
5	X < 40.000	60	12	720

En Mayo de 2002, el tamaño de la muestra de la EPH a nivel Nacional fue de aproximadamente 28500 viviendas.

2.1.7 Rotación de la muestra

Como en cualquier investigación de tipo permanente, en la EPH ha sido necesario utilizar un sistema de renovación del conjunto de unidades a encuestar (del panel de respondentes).

Para lograr esto, las unidades primarias que integran la muestra se agrupan en cuatro "grupos de rotación", cada uno de los cuales es una submuestra de tamaño aproximadamente igual a una cuarta parte de la muestra total.

La rotación se hace de manera que en cada onda, en las unidades primarias pertenecientes al grupo indicado por el esquema de rotación para esa onda, las viviendas seleccionadas anteriormente salen de la muestra para ser reemplazadas por nuevas viviendas seleccionadas dentro de las mismas unidades primarias.

De esta forma, en cada onda permanece en la muestra aproximadamente el 75% de las viviendas encuestadas en la onda anterior.

El sistema elegido permite estimaciones eficientes del cambio entre dos mediciones, evitando la desventaja que significa el aumento de rechazos por encuestar siempre a los mismos hogares.

2.1.8 Actualización de los listados de viviendas

La EPH realiza una actualización periódica previa a cada salida a campo, de los listados de viviendas en las unidades primarias seleccionadas, actualización que permite captar el crecimiento en viviendas dentro de cada una de ellas y mejorar la identificación de las mismas en el terreno. Esta actualización no se realiza en todas las unidades primarias simultáneamente sino por grupo de rotación.

2.1.9 Estimaciones

La expansión de la muestra se realiza en 3 fases :

a) Se expande por la inversa de la probabilidad con que fue seleccionada cada vivienda en la muestra.

b) Se corrige por un *corrector por no respuesta* a nivel de estrato. Para construirlo se utiliza información del cuestionario familiar sobre las causas por las que no pudo realizarse la entrevista. Se clasifican a la causas en dos grupos :

Primer grupo	Segundo grupo
Ausencia	Deshabitada
Rechazo	Demolida
Otras Causas	Fin de Semana
	Construcción
	Vivienda usada solo como establecimiento
	Problemas de listado

El primer grupo corresponde a una no respuesta real, es decir indica viviendas donde *existe* un hogar que no pudo ser contactado. Se define entonces lo que se denominan "viviendas efectivas", que son aquellas que responden, más aquellas que no responden pero presentan causas de no respuesta del primer grupo. Utilizando esta información se construye un corrector Ch :

$$Ch : \frac{\text{Total de viviendas efectivas en el estrato h}}{\text{Total de viviendas que responden en el estrato h}}$$

c) Se realiza un ajuste utilizando las proyecciones demográficas de población elaboradas en INDEC.

2.2 Reformulación muestral de la EPH

La dificultad de dar respuesta a las preguntas relacionadas con la dinámica intraanual de los indicadores del mercado laboral, planteó la necesidad a la EPH, de reformular algunos de sus aspectos muestrales para permitir estudiar mejor su dinámica temporal

Un segundo aspecto que la EPH estudió en distintas oportunidades y que se presentaba como una carencia en la producción estadística es la de no contar con estimaciones para el total nacional, o para el total urbano: Por como se fueron conformando las muestras de la EPH, el dominio del conjunto nacional de las muestras de la EPH es simplemente la agregación de los aglomerados que la componen. No se dispone de una estimación nacional.

Planteados estos dos problemas de la EPH actual, su reformulación se propuso dos objetivos prioritarios desde el punto de vista del diseño muestral:

- *Captación de cambios intraanuales*
- *Ampliación de la cobertura geográfica*

2.2.1 Captación de cambios intraanuales

Para dar cumplimiento a este objetivo se propuso una encuesta de tipo continuo con periodicidad trimestral.

2.2.2 Ampliación de la cobertura geográfica

Respecto del segundo gran objetivo de la reformulación muestral, si bien en un primer momento la EPH planteó un proyecto para simultáneamente pasar a una encuesta continua y ampliar la cobertura geográfica utilizando el Marco de Muestreo Nacional Urbano, la reformulación de dicho Marco y las restricciones presupuestarias indicaron la conveniencia de probar la logística de una encuesta continua en los actuales dominios EPH antes de encarar la ampliación de la cobertura.

Es así como en una primer etapa se abordará el pasaje a una encuesta continua en los actuales aglomerados, para luego en una segunda etapa abordar la extensión geográfica. Se prevé extender la muestra al nuevo Marco Muestral Nacional actualmente en elaboración.

2.3 Principales caracteristicas de la EPH Continua

2.3.1 Dominios de estimación

Trimestralmente, la Encuesta Continua dará estimaciones para los siguientes dominios:

- Total de 28 aglomerados
- Conjunto de aglomerados agrupados por regiones
- Conjunto de aglomerados con 500.000 y más habitantes (según el Censo de oblación 1991)

- Cada uno de los aglomerados con 500.000 y más habitantes (GBA, Gran Rosario, Gran Córdoba, Gran Mendoza, Gran Tucumán, Mar del Plata-Batán, Gran La Plata)
- Resto de los aglomerados EPH

Para no todos los aglomerados el tamaño de muestra permitirá dar estimaciones trimestrales; pero dada la importancia de mantener la serie con los principales indicadores a nivel de aglomerado (serie que lleva ya más de 25 años), se agregarán las muestras trimestrales para dar una *estimación semestral* a nivel de aglomerado de los principales indicadores de la Encuesta.

El tamaño de muestra de los aglomerados con menor población fue ajustado para tener un mínimo de 300 viviendas por trimestre, de forma de obtener en ellos estimaciones semestrales de la tasa de desocupación con un nivel de precisión similar al de las estimaciones actuales. Actualmente el tamaño mínimo de muestra para los aglomerados es de alrededor de 600 viviendas por onda.

Los dominios para los cuales se darán entonces estimaciones a nivel semestral son:

- Total de 28 aglomerados
- Conjunto de aglomerados agrupados por regiones
- Conjunto de aglomerados con 500.000 y más habitantes
- Resto de aglomerados EPH en su conjunto
- Cada uno de los aglomerados EPH

2.3.2 Tamaño de la muestra

Se seleccionará una muestra nacional de 25.000 hogares por trimestre y 100.000 hogares por año [1], bajo los siguientes supuestos:

Tasa de desempleo: 14%

Coeficiente de correlación para la variable desocupación entre dos ondas consecutivas: 0.2

Efecto diseño: 1.5

Promedio de personas económicamente activas por hogar: 1.5

Nivel de confianza: 90%

Este tamaño de muestra permitirá detectar diferencias significativas de al menos un 0.5% entre dos estimaciones de la tasa de desempleo de dos trimestres consecutivos a nivel nacional y de por lo menos un 1% a nivel regional.

Bajo estas condiciones se logrará estimar la tasa nacional de desempleo con un coeficiente de variación inferior al 2% y las tasas por región con un coeficiente de variación del 5%.

2.3.3 Asignación de la muestra

Los 25000 hogares por trimestre se distribuyeron en:

Aglomerados EPH : 17000

Resto : 8000

A su vez dentro de los aglomerados EPH se asignó la muestra tratando de aproximarse lo más posible a una muestra autoponderada.

[1] Ver documento presentado en la Primer Reunión sobre Estadística Púbica del Instituto Inteamericano de Estadística, Buenos Aires, Junio de 1998.

Las restricciones que hubo que respetar fueron: i) Un tamaño mínimo de muestra que permitiese, para los aglomerados mas pequeños, mediante la agregación de muestra, dar estimaciones semestrales con niveles de precisión similares a los actuales. ii) Igual cantidad de viviendas a seleccionar por UPE (por razones logísticas).

2.3.4 Diseño de la muestra

Dentro de cada aglomerado EPH se mantuvieron las actuales UPEs distribuyéndolas a lo largo de las 12 semanas del trimestre. . Esta distribución es fija, o sea que una UPE es encuestada siempre en la misma semana de cada trimestre.

El tamaño de los conglomerados, o sea la cantidad de unidades de segunda etapa a seleccionar en cada UPE es constante y difiere levemente entre aglomerados. Ver Anexo B con los tamaños de muestra propuestos.

En aquellos donde el tamaño de muestra de la encuesta continua es menor al actual se decidió mantener la cantidad de UPEs reduciendo las viviendas seleccionadas dentro de ellas. Esto presenta la ventaja de mantener la cantidad de "puntos muestra", de forma de minimizar el aumento de la varianza. Y en caso de desearse una muestra mayor basta con aumentar la muestra seleccionanda dentro de las UPEs.

2.3.5 Asignación de las UPEs a las semanas del trimestre

La distribución se hizo de modo que asegurara :

- Distribución homogénea de los estratos a lo largo del trimestre
- Igual cantidad de UPEs a relevar a nivel nacional por semana
- Similar cantidad de UPEs por mes en cada aglomerado
- Similar cantidad de UPEs por mes y grupo de rotación

Ver Anexo A con la distribución de la muestra

2.3.6 Esquema de rotación

Un aspecto importante a definir en función de los objetivos de la Encuesta es el esquema de rotación a adoptar.

El esquema de rotación de una encuesta que se repite en el tiempo tiene incidencia sobre cuatro aspectos:

a) Precisión de las estimaciones del cambio entre dos períodos consecutivos, o entre iguales períodos de años consecutivos, etc.

b) Precisión de las estimaciones obtenidas al agregar muestra, por ejemplo en una encuesta trimestral al agregar muestra para obtener estimaciones semestrales.

c) El aumento de la no respuesta por el cansancio del panel.

d) La posibilidad de realizar estudios longitudinales siguiendo a un hogar o a un respondente desde su ingreso en el panel hasta la última vez que es encuestado.

El solapamiento de las muestras entre dos períodos consecutivos (o sea el porcentaje de muestra en común) juega en sentido contrario para los primeros dos aspectos: Si un esquema tiene un alto porcentaje de solapamiento entre un período y el siguiente, medirá bien los cambios pero disminuirá su precisión para una agregación a lo largo de varios períodos. Por el contrario, un bajo solapamiento no mejora sensiblemente la captación de los cambios, pero posibilita agregar muestra a lo largo de varios períodos, para dar estimaciones para dominios mas reducidos.

Atendiendo estos argumentos se optó por el esquema (2-2-2), ya probado en la prueba piloto realizada en el aglomerado GBA, teniendo en cuenta para ello la experiencia previa, en lo que respecta al efecto en la tasa de no respuesta de largos periodos de permanencia en el panel, la necesidad de contar con un solapamiento mínimo entre iguales períodos de años consecutivos y de períodos consecutivos. Se tuvo en cuenta también, la necesidad de muchos usuarios consultados, de poder realizar estudios de panel con las muestras de la EPH. En otros países con más recursos es común la realización de estudios longitudinales, pero dado lo costoso de los mismos, la EPH consideró importante que los investigadores contaran con un esquema que les permitiese realizar un seguimiento de los hogares, aunque fuera por períodos cortos de tiempo. En el esquema de rotación elegido, las viviendas permanecen dos trimestres en la muestra, descansan dos trimestres y vuelven otros dos trimestres a la misma hasta finalmente abandonarla.

Este esquema asegura un solapamiento de 50 % entre trimestres consecutivos y entre los mismos trimestres entre años consecutivos.

Anexo A

Distribucion de areas por semana y grupo de rotacion

		Grupo de rotación				Total
		1	2	3	4	
Semana	1	49	48	56	45	198
	2	50	50	46	52	198
	3	44	56	49	48	197
	4	56	46	46	50	198
	5	49	49	50	50	198
	6	51	47	53	47	198
	7	51	52	45	50	198
	8	49	49	48	51	197
	9	48	49	50	51	198
	10	53	54	47	43	197
	11	55	39	51	53	198
	12	38	54	52	53	197
Total		593	593	593	593	2372

Distribucion de areas por mes y grupo de rotacion

		Grupo de rotación				Total
		1	2	3	4	
Mes	1	199	200	197	195	791
	2	200	197	196	198	791
	3	194	196	200	200	790
Total		593	593	593	593	2372

Viviendas seleccionadas por semana y grupo de rotación

		Grupo de rotación				Total
		1	2	3	4	
		Tot viv	Tot viv	Tot viv	Tot viv	Tot viv
Semana	1	360	351	414	312	1437
	2	356	364	334	383	1437
	3	321	408	362	344	1435
	4	410	332	322	374	1438
	5	362	356	357	359	1434
	6	367	334	396	345	1442
	7	371	382	323	367	1443
	8	358	356	350	369	1433
	9	357	353	354	374	1438
	10	365	393	352	324	1434
	11	397	297	363	386	1443
	12	286	384	383	373	1426
Total		4310	4310	4310	4310	17240

Viviendas seleccionadas por mes y grupo de rotación

		Grupo de rotación				Total
		1	2	3	4	
		Tot viv	Tot viv	Tot viv	Tot viv	Tot viv
Mes	1	1447	1455	1432	1413	5747
	2	1458	1428	1426	1440	5752
	3	1405	1427	1452	1457	5741
Total		4310	4310	4310	4310	17240

Anexo B

EPH Continua - Muestra asignada a los aglomerados actuales

Aglomerados	Muestra en el trimestre	Total de areas	Viviendas por area
Río Gallegos	288	48	6
Ushuaia-Rio Grande	300	60	5
Santa Rosa-Toay	300	60	5
La Rioja	300	60	5
Concordia	300	60	5
Gran Catamarca	300	60	5
San Luis-El Chorrillo	300	60	5
Cdoro. Rivadavia-R.Tilly	300	60	5
Río Cuarto	300	60	5
Formosa	300	60	5
Jujuy-Palpalá	420	60	7
Neuquén-Plottier	300	60	5
Sgo. del Estero-La Banda	420	60	7
Gran Paraná	420	60	7
Corrientes	420	60	7
Posadas	480	60	8
Gran Resistencia	480	60	8
Salta	480	80	6
Gran San Juan	480	80	6
Bahía Blanca-Cerri	480	80	6
Gran Santa Fe	480	80	6
GranTucumán-Tafi Viejo	800	80	10
Gran Mendoza	800	80	10
Gran La Plata	800	80	10
Mar del Plata-Batán	800	80	10
Gran Rosario	1000	100	10
Gran Córdoba	1000	100	10
Total interior	**13048**	**1848**	
GBA	4192	524	8
Total 28 aglomerados	**17240**	**2372**	

INFORME TÉCNICO DEL DISEÑO DE LA MUESTRA MAESTRA DE HOGARES

FERNANDO RIVERO Y FERNANDO MOLLINEDO

INSTITUTO NACIONAL DE ESTADÍSTICA (INE)
BOLIVIA

ÍNDICE

Página

1.- Introducción y Objetivos de la Muestra Maestra 119
2.- Actividades a Desarrollarse 119
3.- ¿Qué es una Muestra Maestra (MM) para Bolivia? 119
4.- ¿Por Qué una MM? 120
5.- Construcción de la Muestra Maestra 120
6.- Tipo de Muestreo 121
7.- Tareas para la Muestra Maestra 121
8.- Las Unidades Primarias de Muestreo (UPMs) 122
9.- Selección de las UPM's 122
10.- Construcción de Estratos 122
11. El Tamaño de la MM 123
12. El Marco para la Muestra Maestra 123
13.- Afijación de las UPMs 123
14. Actualización de la Muestra Maestra 124
15. Selección de una Nueva Muestra Maestra 124
16. Estimación y Factores de Expansión 124
17. Muestras de Rotación, Paneles 125
18. Selección de las Muestras y Asignación a Paneles 126
19. Unidades de Observación que no Permanecen en la Muestra 127

UNA MUESTRA MAESTRA PARA BOLIVIA

1. Introducción y objetivo de la Muestra Maestra

El Instituto Nacional de Estadística de Bolivia (INE) ha realizado un Censo de Población y Vivienda en Septiembre de 2001. A partir de la información recolectada por este censo se pretende construir una Muestra Maestra para todas las encuestas de hogares que el INE tiene previsto realizar.

Objetivo Principal

El Objetivo principal es proponer una metodología para la construcción de la Muestra Maestra, que sirva de marco para todas las encuestas de hogares que tiene previsto realizar el INE en los próximos 10 años, dividido en dos partes de 5 años cada uno, teniendo como insumo principal la información cartográfica y la base de datos del Censo de Población y Vivienda de 2001.

2. Actividades a Desarrollarse

Definido el objetivo principal, se plantea las siguientes actividades:

a. Desarrollar la metodología para la construcción de la Muestra Maestra.

b. Preparar y construir el Marco Muestral para la selección de dicha muestra.

c. Conformar la metodología de estimación para las encuestas continuas bajo un esquema de rotación de muestras que sea sostenible en el tiempo.

d. Proponer formas de actualización de la Muestra Maestra, en relación a las áreas de nueva creación de viviendas.

3. ¿Qué es una Muestra Maestra (MM) para Bolivia?

Una Muestra Maestra en términos generales es: definir un tamaño de muestra, plantear un tipo de muestreo, una forma de selección y finalmente definir una metodología de estimación de parámetros y de sus varianzas, a partir de seleccionar una muestra de la cual se obtienen sub muestras para cada una de las encuestas a realizarse, durante un periodo de tiempo.

Por tanto no se seleccionan muestras independientes para cada encuesta, como se hace cotidianamente, cuando no se tiene una muestra maestra.

La Muestra Maestra de Bolivia, es una muestra de Unidades Primarias de Muestreo (UPMs), donde las UPMs son áreas geográficas que contienen entre 150 y 300 viviendas ocupadas. La división de las áreas geográficas, respetan los límites de las divisiones político administrativo del país y los límites censales de sector y segmento. Además, es importante definir dicha delimitación en base a las *Organizaciones Comunitarias* que tiene el área rural dispersa y que en su mayoría son reconocidas como tal por los pobladores del área para evitar problemas de mala delimitación. Las UPMs estarán constituidas por una Organización Comunitaria, un grupo de organizaciones comunitarias o de una parte de ella. En el área urbana, una posibilidad es el Sector Censal que es una agrupación de manzanos y que contiene en su interior en promedio 120 viviendas. El tamaño de la UPM, está en función a la cantidad de encuestas que el INE tiene planificado ejecutar el los próximos cinco años, y del diseño de la muestra planteado.

Definidas y seleccionadas las Unidades Primarias de Muestreo de la MM, se seleccionarán las Unidades Secundarias de Muestreo (USMs) que en nuestro caso serán segmentos censales, manzanos o directamente viviendas.

En cada UPM seleccionada de la Muestra Maestra, se efectuará un listado de una parte de la misma, para contar con un dato actualizado en el momento del operativo de la encuesta. Donde este listado viene de la

aplicación del Formulario 1, implementado en el censo 2001, donde se recoge información del jefe de hogar.

Para la selección de USMs se pueden aplicar criterios de estratificación diferentes a los de las UPMs; por ejemplo el nivel socio-económico (pobreza). En esta selección, es recomendable introducir un período de descanso para los hogares que participaron ya en una encuesta.

4. ¿Por qué una MM?

Una Muestra Maestra concentra la preparación de muestras y el trabajo de campo en un número limitado de UPMs. La ventaja más importante, es la *reducción de costos, el reconocimiento permanente del área muestreada y la reducción de los listados de actualización.* Las mismas UPMs (o una parte de ellas) se usan para todas las encuestas y un listado se puede normalmente considerar actual por un año aproximadamente.

Otra ventaja es la *reducción del trabajo de muestreo.* En vez de un marco de todos los sectores y segmentos del Censo, basta un marco de las UPMs en la muestra.

Otra ventaja es la correlación de la información de varias encuestas lo que facilita la complementariedad de temas de investigación evitando la duplicidad de esfuerzos.

El número de viviendas ocupadas en Bolivia, es de 2.2 millones; la Muestra Maestra propone seleccionar alrededor de 125.000 viviendas, que hacen aproximadamente una cantidad de 1.200 UPM's, para 5 años de trabajo de encuestas, esta muestra es suficiente para cubrir todas las encuestas que el INE tiene planificado realizar.

La desagregación de la Muestra Maestra se describe en el siguiente cuadro:

Distribución de la Muestra Maestra (viviendas)

Año 2002	Año 2003	Año 2004	Año 2005	Año 2006	Total
15.000	50.000	20.000	20.000	20.000	125.000

Para el año 2003, la muestra es 50.000 viviendas, ya que se tiene prevista una Encuesta de Mortalidad Materna introducida en la Encuesta Nacional de Demografía y Salud (ENDSA), con desagregaciones a nivel departamental.

Una desventaja eventual es que la MM es *una muestra general.* Ella no puede ser óptima en cada encuesta, particularmente se precisará una estratificación especial. Pero la MM es muy flexible, y se adapta a exigencias diversas.

Para el cálculo de los estimadores, se plantea introducir estimadores de regresión, que usa información auxiliar o adicional de las mismas UPMs, a partir de la información censal.

Si bien los estimadores de regresión no son estimadores insesgados, generalmente estos estimadores presentan una precisión muy elevada que otros estimadores por la utilización de la información auxiliar.

5. Construcción de la Muestra Maestra

Insumos

- La cartografía de la Actualización del Censo del 2001, corregida con los datos censales finales.
- La base de datos Censal desagregados tanto a nivel de vivienda como de población, para luego hacer las agrupaciones pertinentes para la construcción de UPMs .

Para la construcción de las UPM's, todas las operaciones son realizadas en base a la información proveniente del Censo, específicamente la información de las viviendas y de la población en cada una de ellas.

La información de la Actualización Cartográfica del Censo es la información que permite fijar los límites geográficos de las UPM's en términos de la cantidad de viviendas ocupadas que tiene cada una de ellas. También permite medir posibles cambios entre la actualización y el censo.

6. Tipo de Muestreo

Se tiene previsto un diseño de muestra de tipo trietápico para la Muestra Maestra. Es decir que:

Las Unidades de Primera Etapa (UPM's), en su generalidad, son las Organizaciones Comunitarias para el área rural dispersa, sin embargo en algunos casos se deben realizar agregaciones o particiones respetando límites de cantón, sector y segmento censal para contar con cierto número de viviendas en cada UPM . Los manzanos establecidos en un sector censal, son las UPM's para el área urbana, de igual forma, será importante contar con cierta cantidad de viviendas que respondan a los requerimientos de la Muestra Maestra. La selección de dichas UPM's están en base a la probabilidad proporcional al tamaño de viviendas que cuente.

Las Unidades de Segunda Etapa serán los segmentos censales para el área dispersa y el manzano para el área urbana. Su selección está en base a un muestreo aleatorio simple con igual probabilidad ya que presentan una estandarización en el número de viviendas que tiene.

Y finalmente las Unidades Ultimas de Muestreo son las viviendas con ocupantes presentes. Su selección es sistemática con arranque aleatorio.

La definición del número optimo de viviendas por UPM se obtendrá una vez consolidada la base censal y la construcción del marco muestral. Donde se precisarán la cantidad de UPM's y la cantidad de viviendas a encuestarse en cada UPM, distribuidas en todo el país de acuerdo a los resultados del análisis de correlación intraclásico que se haga.

7. Tareas para la Muestra Maestra

La MM debe servir para las siguientes encuestas y otras que se actualizarán. El detalle viene en el siguiente cuadro:

Año	*Encuesta*	*Desagregación de los resultados*	*Número previsto de viviendas*
2002	Encuesta de Hogar MECOVI Ampliada (con modulo empleo)	Departamento, Urbana/rural, 3 pisos ecológicos	12.000
2002	Demografía y Salud (ENDSA)	Departamento, Urbana/rural, 3 pisos ecológicos, 5 niveles de pobreza	30.000
2003-	- Encuesta de Hogares MECOVI	Urbana/rural, 3 pisos ecológicos	6.000
2004-	- Encuesta de Hogares MECOVI	Urbana/rural, 3 pisos ecológicos	6.000
2005	- Encuesta de Hogares MECOVI	Urbana/rural, 3 pisos ecológicos	6.000
2006	- Encuesta de Hogares MECOVI	Urbana/rural, 3 pisos ecológicos	6.000

Estas son las encuestas planificadas para los próximos 5 años, y para este número de encuestas se tiene una muestra de 125.000 viviendas aproximadamente.

De haber otros requerimiento como las encuestas municipales, podrían determinar un tamaño distinto para la MM.

8. Las Unidades Primarias de Muestreo (UPMs)

En general, las unidades primarias de muestreo deben reflejar la variación en cada estrato (de las variables de las encuestas). La agrupación de las UPM's deben contener correlaciones intraclásicas pequeñas que pueda aumentar la precisión de las estimaciones. Es por ese motivo, se hace una diferenciación entre áreas geográficas, tomando como delimitaciones en el área urbana los Sectores Censales y las Organización Comunitarias para el área dispersa.

También permitirá identificar las UPM's fácilmente en el terreno, y que además formen áreas conectadas preferiblemente compactas, para optimizar el uso de los recursos económicos y otras actividades relacionadas al trabajo del campo.

9. Selección de las UPM's

La selección en la primera etapa es proporcional al tamaño de viviendas que tenga (ppt), como anteriormente se había dicho. Las UPM's conformadas en el marco de muestreo tienden a ser Organizaciones Comunitarias en su mayoría lo que genera una dispersión en la cantidad de viviendas entre 150 y 300.

En algunos casos, las UPM's están constituidas por un grupo de Organizaciones Comunitarias que juntas hacen entre el rango de viviendas establecido anteriormente y lo mismo si se trata de una organización comunitaria demasiado grande a la cual se la debe dividir por el mismo concepto. En el caso de una división, es importante respetar los límites de cantón, sector y segmento censal.

10. Construcción de Estratos

Los estratos para la MM deben seleccionarse considerando, en parte, las desagregaciones deseadas en las encuestas. Es preferible crear estratos correspondientes a las subdivisiones geográficas del país.

Para Bolivia las estratificaciones mas importantes son:

Area Urbano y Rural.

Regional Altiplano Valle y Llano.

Departamentos

Las desagregaciones anteriores también son utilizadas como Dominios de Estudio.

Otras posibilidades de estratificación se debe a la desagregación de población :

Ciudades capitales

Otras ciudades con 10.000 habitantes o más

Ciudades con 2.000 a 10.000 habitantes

Ciudades y aglomeraciones con 250 - 2.000 habitantes

Otras áreas denominadas áreas dispersas.

Los tres primeros estratos son considerados urbanos, los otros rurales.

También se puede estratificar por categorías de pobreza u otros indicadores socio-económicos, para lo cual se estudiará la posibilidad de emplear los métodos univariados de Dalenius y Hodges o para métodos multivariados mediante Cluster Análisis.

11. El tamaño de la MM

El tamaño de la Muestra Maestra, está en relación con las encuestas que el INE tiene previsto ejecutar en los próximos cinco años. En base a esto, se ha fijado una muestra de 125.000 viviendas ocupadas.

Si se plantea una cantidad de 50 a 60 viviendas por unidad primaria de muestreo (USM), se tendrá aproximadamente un total de 2.500 UPM's.

Una alternativa a definir, es el cupo de encuestas por USM, que puede establecerse entre 8 a 10 viviendas en el área amanzanada, y 12 a 16 viviendas en el área dispersa. Esto dependerá de cuán dispersa se quiere la muestra o cuán agregada debe ser. Esto está también en función de costos y tiempo.

12. El Marco para la Muestra Maestra

El Censo del 2001 proporciona datos a nivel:

Departamento

Provincia

Sección municipal

Cantón

Localidad

Organización Comunitaria

Zona

Sector

Segmento

Manzano

Descripción de área urbano / rural.

Tipo de vivienda

Número de personas por vivienda

A partir de esta información se construirá el Marco Muestral para las encuestas de hogares y para la selección de la Muestra Maestra.

La descripción de la Unidades de Muestreo que ya han sido mencionadas anteriormente, serán incorporadas en el marco elaborado. De igual manera, las UPM's serán incorparadas en la cartografía.

A la fecha, el trabajo de la delimitación de UPM's está en plena ejecución.

Cabe mencionar que para producir este marco, se precisan *datos definitivos del censo* de todos los sectores y segmentos censales, con sus correspondientes identificadores, número de viviendas y los mapas ya consolidados que compatibilicen entre la cartográfica y la base de datos censal.

La lista de las Unidades de Muestreo con la descripción de estratos y el número de viviendas en cada uno, forman el Marco Muestral para la selección de la Muestra Maestra.

13. Afijación de las UPMs

Se propone un método de afijación potencial, con un número de UPMs en cada estrato proporcional a M_h^{α}, donde M_h = el número de viviendas en el estrato h, y α es un parámetro a seleccionar – con $\alpha = 0.5$. Este método ha sido implementado en varias encuestas de hogares en el país, con resultados satisfactorios.

La afijación potencial aumenta el tamaño de la muestra en los estratos con pocas viviendas, lo que produce buenas estimaciones en los estratos pequeños sin perder la precisión en los estratos grandes.

Sin embargo no está demás discutir otro método de afijación que mejore las estimaciones.

14. Actualización de la Muestra Maestra

Para la actualización de la Muestra Maestra, se empleará el método de listado de actualización de las Unidades de Segunda Etapa seleccionadas en primera instancia, hasta que esta sea agotada con las primeras encuestas.

Posteriormente se seleccionará una nueva USM y así sucesivamente hasta tener cubierta toda la UPM seleccionada y que su actualización sea total.

Los listados de las UPMs darán nuevos totales de viviendas. Los nuevos resultados se anotan en el registro y se usan para recalcular los factores de expansión. El uso de listados en una MM no es diferente a una muestra tradicional.

15. Selección de una Nueva Muestra Maestra

La primera Muestra Maestra no se incluye para la selección de una segunda, en esta situación se crea una nueva Muestra Maestra independiente de la primera, y esta se basa en una actualización del total del marco muestral. Para esto se utilizan los listados y resultados de la primera muestra y la posibilidad de realizar una nueva actualización cartográfica.

En el caso de no poder lograr una nueva actualización cartográfica, se propone actualizar solamente las zonas de crecimiento, especialmente del área urbana. Esto dependerá de los recursos económicos disponibles para levantar un trabajo operativo de actualización.

16. Estimación y Factores de Expansión

La estimación en una encuesta que utiliza una Muestra Maestra es similar a las estimaciones de otras encuestas, donde los factores de expansión requieren de una atención especial.

Un total T se estima por

$$\hat{T} = \sum_{hji} w_{hji} x_{hji}$$

donde

x_{hji} = valor observado en el hogar i (o la persona i) en la UPM j del estrato h

w_{hji} = factor de expansión de la observación i en la UPM j del estrato h (ver abajo),

y la suma se extiende sobre los estratos h relevantes, incluyendo todas las UPMs y todos los hogares (personas) en estos estratos.

El promedio se obtiene en la manera usual, dividiendo con $\hat{N} = \Sigma\, w_{hji}$.

La tasas se obtienen como promedios, con $x_{hji} = 0$ o 1.

En el cálculo de la varianza de un promedio o una tasa, el estimador es una razón de dos cantidades aleatorias, ya que el numerador $\hat{N}$ es un estimador también. La fórmula para la varianza aproximada de una razón $\hat{R} = \hat{T} / \hat{N}$ es

$$Var(\hat{R}) = Var(\hat{T}) + \hat{R}^2 Var(\hat{N}) - 2\hat{R}Cov(\hat{T}, \hat{N})$$

Definiendo $x_{hj+} = \sum_{i=1}^{n_{hj}} w_{hji} x_{hji}$, la varianza de $\hat{T} = \sum_{hji} w_{hji} x_{hji}$ se puede estimar con

$$Var(\hat{T}) = \sum_{h=1}^{H} \frac{n_h}{n_h - 1} \left[\sum_{j=1}^{n_h} x_{hj+}^2 - (\sum_{j=1}^{n_h} x_{hj+})^2 / n_h \right]$$

donde H = el número de estratos, y n_h = el número de UPMs seleccionadas en el estrato h.

La covarianza se estima de una manera análoga:

$$Cov(\hat{T}, \hat{Y}) = \sum_{h=1}^{H} \frac{n_h}{n_h - 1} \left[\sum_{j=1}^{n_h} x_{hj+} y_{hj+} - (\sum_{j=1}^{n_h} x_{hj+})(\sum_{j=1}^{n_h} y_{hj+}) / n_h \right]$$

donde "y" es otra variable, $\hat{Y}$ es el total estimado y y_{hj+} se define como x_{hj+}.

Las estimaciones y varianzas pueden ser calculadas mediante los programas Pccarp, Cluster o Stata.

Si la selección de USMs es aleatoria, el factor de expansión es

(F1) $$w_{hji} = \frac{N_h}{n_h} \frac{N'_{hj}}{N_{hj}} \frac{}{n_{hj}},$$

donde

N_h = número de viviendas en el estrato h

n_h = número de UPMs seleccionadas (con ppt) en el estrato h

N_{hj} = número de viviendas en la UPM j del estrato h si fue seleccionada la UPM

N'_{hj} = número de viviendas al último listado en la UPM j del estrato h

n_{hj} = número de viviendas seleccionadas para la muestra en la UPM j del estrato h

En el factor de expansión se puede incluir también:

- Una corrección de no-respuesta - a menudo el factor n_{hj}/n^*_{hj}, donde n^*_{hj} es el número de respuestas obtenidas
- Una corrección, o calibración, para el cambio (crecimiento) en la población total de viviendas en el estrato - usando el factor N^*_h/N_h, donde N^*_h es la población total estimada o proyectada al tiempo de la encuesta

En el área dispersa, donde se trabaja con segmentos (componentes), basta utilizar N_{hj} y N'_{hj} como poblaciones del segmento.

Si la UPM es seleccionada (con PPT) en un estrato, en la actualización de la MM, todos los factores de expansión en el estrato cambian. La fórmula (F1) se aplica, con n_h igual a todas las UPMs seleccionadas (entonces, n_h nuevo = n_h viejo + b).

17. Muestras con Rotación, Paneles

Una posibilidad para muestras con rotación son usadas cuando una encuesta se repite cada año (o cada trimestre, cada mes) y cuando hay un interés particular de estimar los cambios entre las encuestas. Cada vivienda permanece en la muestra un determinado número de rondas, sea 2, 3, 4 o más. Una fracción de la

muestra sale cada vez y el mismo número de viviendas entran. La fracción corresponde al número de rondas de permanencia - con dos rondas de permanencia, ½ de la muestra sale cada vez, con 3 rondas, 1/3 sale etc. La tasa de rotación en estos casos es 50%, 33%, etc.

En la situación de las encuestas de hogares en Bolivia, sería más natural realizar la rotación al nivel de las UPMs. En cada ronda, saldría un grupo de UPMs de la muestra, y otro grupo entraría.

En cada UPM, las viviendas participantes son las mismas durante todo el período que la UPM permanece en la muestra.

Un grupo de UPMs, y las viviendas seleccionadas en aquellas, que permanecen durante las mismas rondas se puede llamar un panel. Por ejemplo, con una tasa de rotación de 25 % (o sea 4 rondas de participación), se puede representar las muestras de las primeras 5 rondas así:

Muestra de la primera ronda	P1 P2 P3 P4	
Muestra de la segunda ronda	P2 P3 P4 P5	(el panel P1 sale, y P5 entra)
Muestra de la tercera ronda	P3 P4 P5 P6	(el panel P2 sale, y P6 entra)
Muestra de la cuarta ronda	P4 P5 P6 P7	
Muestra de la quinta ronda	P5 P6 P7 P8	(aquí, no queda ningún panel de la primera ronda)

Como un caso extremo, es posible trabajar con una encuesta de un panel fijo, con una tasa de rotación de 0 %. No es recomendable para encuestas de hogares, porque un panel fijo pierde muy rápidamente la representatividad.

Las ventajas más importantes de las muestras de rotación son:

a) Posibilidades de estimar cambios (entre rondas de la encuesta) con mayor precisión

b) Reducción del trabajo del campo, ya sea en listados, o en buscar los hogares o en solicitar las respuestas - las personas que participan conocen las preguntas (y los encuestadores) desde la primera vez

c) La boleta o el número de preguntas de la segunda ronda y rondas posteriores puede ser reducida

Algunas desventajas son:

a) La estimación de agregados (por ejemplo un promedio o un total anual) es menos precisa

b) Los participantes se cansan y no permanecen en la muestra todo el tiempo previsto, lo que significa que se pierde una parte de las ventajas

c) Deben ser establecidos procedimientos para salidas no planificadas de la muestra, así como para entradas en la muestra

El parámetro más importante a determinar es la fracción de rotación, o el número de rondas de participación de cada hogar. Hay dos consideraciones contrarias. Una es la importancia de las estimaciones de cambio versus las de niveles - si los cambios son muy importantes, una gran parte de los hogares deben permanecer en la muestra, es decir cada hogar debe participar muchas veces. Otra es la atención a los hogares, y el riesgo de perder hogares en entrevistas repetidas. Esto resulta en un número bajo de rondas de participación.

18. Selección de las muestras y asignación a paneles

La primera muestra se selecciona como cualquier muestra. Normalmente se selecciona las UPMs con ppt (viviendas). Si hay una MM, y si precisa una selección de la misma, esta selección debe ser sistemática o

aleatoria. Asumiendo que queremos una tasa de rotación de 25 %, la muestra debe ser dividida en cuatro paneles. Por eso, cada UPM seleccionada será asignada a uno de los paneles P1,..., P4. Esto se realiza con una lista de las UPMs seleccionadas en orden de los estratos. Empezando con un panel aleatorio (por ejemplo P3), la primera UPM de la lista es asignada a P3, la segunda UPM a P4, la tercera a P1, etc. En un estrato nuevo, se continúa con la secuencia. El número del panel debe ser incluido en la lista de las UPMs de la muestra.

Para la selección de la segunda muestra, precisa sólo el panel P5, que debe sustituir al panel P1. Los otros paneles, con todas las UPMs y viviendas que contengan, permanecen en la muestra. El número de UPMs en P5 y la distribución por estratos es completamente determinado por P1. Basta seleccionar el mismo número de UPMs, en los mismos estratos. Si hay una MM, las UPMs se seleccionan de la MM como antes. Si no hay una MM, la selección debe ser con ppt.

Se propone que la reselección de una UPM que sale con P1, sea permisible inmediatamente, mas que las viviendas a seleccionar en la UPM puedan gozar de un período de descanso de 2 años. Por lo contrario, no se permite que una UPM entre en dos o más paneles que son usados simultáneamente. Entonces, en la selección de P5, las UPMs de P2-P4 no son permisibles.

En una encuesta trimestral, será conveniente seleccionar todos los paneles necesarios durante un año, a decir cuatro paneles, simultáneamente en una sola operación. En este caso se propone que todas las UPMs seleccionadas sean diferentes (sin reposición). Con esta regla, la reselección de una UPM será posible solo después de un año.

19. Unidades de observación que no permanecen en la muestra

Por varias razones, hogares y personas no permanecerán en la muestra por todo el período previsto. Sin experiencia de la probable frecuencia, es necesario seguir el desenvolvimiento con mucha atención, particularmente en la segunda ronda de la encuesta. En principio, se prefiere trabajar sin substituciones. Si el problema es muy grande, precisa entonces pensar en términos de un aumento de la muestra bruta (o posiblemente todavía una substitución).

Si un hogar transfiere a otro lugar, o a cualquier vivienda no incluida en la muestra, se propone no seguir más este hogar. Por lo contrario, si un nuevo hogar entrara en una vivienda de la muestra, se propone que ello sea incluido en la muestra. Se debe evaluar el efecto de esta regla después de la segunda y tercera ronda, para decidir definitivamente si se quiere o no conservarla.

El efecto de una muestra de rotación es reducir el error estándar de diferencias, dado que se crea una correlación entre los valores de dos años. Los promedios de dos o tres años pueden ser indicadores interesantes para desagregación en áreas como los departamentos, donde el número de observaciones en un año no es suficiente para estimaciones con precisión deseada. Estos promedios sufren de la muestra de rotación, en cuanto aumenta el error estándar.

Se nota que es necesario hacer suposiciones del tamaño de las correlaciones ρ_1 y ρ_2 para las variables diferentes. No hay ninguna información, pero es probable que sean altas, y en los cálculos son supuestos valores de 0,8 - 0,95 para ρ_1 y 0,6 - 0,9 para ρ_2.

Se nota que el error estándar para la diferencia de dos rondas consecutivas es siempre igual al error en una ronda, multiplicado por $\sqrt{(2-2\beta_1\rho_1)}$. Si tr = la tasa de rotación, $\beta_1 = 1 - tr$, y β_1 es la proporción de la muestra que permanece en las dos rondas consecutivas. (Además, tr = 100% (o $\beta_1 = 0$) corresponde al caso de muestras independientes, todos los hogares salen de la muestra después de una ronda.) Es interesante estudiar la tasa de reducción del error estándar en relación al caso de muestras independientes. Si D denota el error estándar, esta tasa es

$$\frac{D_{con\,rot}(t_{i+1}-t_i)}{D_{\sin rot}(t_{i+1}-t_i)}=\frac{\sqrt{Var(t)}\sqrt{(2-2\beta_1\rho_1)}}{\sqrt{Var(t)}\sqrt{2}}=\sqrt{1-\beta_1\rho_1}\,.$$

El error estándar de una diferencia, con muestras de rotación, es 60 - 70 % del error para muestras independientes. Si la correlación ρ_1 es muy alta, la diferencia de tasas diferentes es importante.

Por el contrario, cuando se consideran promedios de dos o tres años, se pierde de precisión con muestras de rotación. La razón de los errores estándar

$$\frac{D_{conrot}((t_{i+1}+t_i)/2)}{D_{\sin rot}((t_{i+1}+t_i)/2)}=\frac{\sqrt{Var(t)}\sqrt{(1+\beta_1\rho_1)/2}}{\sqrt{Var(t)/2}}=\sqrt{1+\beta_1\rho_1}$$

Si el error estándar aumenta de 20 a 30 %. Las alternativas (combinaciones de Tasa y ρ_1) no son muy diferentes.

Para un promedio de tres años, precisan dos correlaciones. La razón relevante es

$$\sqrt{(3+4\beta_1\rho_1+2\beta_2\rho_2)/9}\,/\,\sqrt{1/3}\,.$$

INFORME TÉCNICO SOBRE DISEÑO DE UN MARCO DE MUESTREO DE ÁREAS PARA LAS ENCUESTAS DE HOGARES

INSTITUTO NACIONAL DE ESTADÍSTICA (INE)
CHILE

ÍNDICE

Página

Introducción 131
1.- Censo y Marco de Muestreo de Áreas 131
2.- Construcción del Marco de Muestreo de Áreas 133
2.1 Formación de áreas 133
2.2 Conformación de estratos del marco de muestreo 133
2.3 Unidades de primera etapa (UPE) o secciones 134
2.4 Tamaño óptimo de las unidades de primera etapa o secciones 134
2.5 Actualización de la áreas seleccionadas 134
2.6 Selección de una muestra en etapas 135
3.- Descripción del Diseño Muestral Utilizado para la Encuesta Nacional del Empleo 135
3.1 Población objetivo 135
3.2 Cobertura geográfica y áreas de estimación 135
3.3 Unidades de muestreo y unidades de información 136
3.4 Determinación del tamaño de la muestra 136
3.5 Selección de la muestra 137
3.6 Distribución de la muestra 137
3.7 Actualización y enumeración de secciones para rotación de viviendas seleccionadas 137
4.- Avance De la Actualización del Marco de Muestreo de Áreas con Precenso 2001 138
Referencias 140
Anexo 141

Introducción

En el año 1985, con los antecedentes del Censo de Población y Vivienda de 1982, se construyó un nuevo marco llamado "Programa Integrado de Encuestas de Hogares" (PIDEH), dado los nuevos antecedentes aportados por el Precenso de 1991 y el Censo de Población y vivienda de 1992 se actualizo este marco y se selecciono una muestra de 34269 viviendas particulares sobre la cual se aplica en forma continua la Encuesta Nacional del Empleo (ENE), que rigen actualmente.

Las ventajas de un marco de muestreo de áreas para eventuales estudio sobre población y vivienda son evidentes. Las muestras seleccionadas por áreas son más económicas y en el caso de una encuesta continua es posible actualizar o renovar la muestra sin quebrar las reglas de una selección probabilística.

El Marco Muestral de áreas esta dividido en 25826 secciones o áreas de 200 viviendas para las áreas urbanas y 100 en las áreas rurales. Este marco se presta especialmente para la selección de muestras de viviendas en dos etapas a nivel Nacional, regional, provincial y grandes centros urbanos.

Este informe se divide en tres partes. La primera parte corresponde a la información obtenida del Censo para la construcción del marco. La segunda parte se refiere a la construcción del marco de muestreo de áreas respecto a formación de áreas, de estratos, tamaño optimo de las unidades de primera etapa. La tercera parte describe el diseño muestral utilizado para la Encuesta Nacional del Empleo.

Cabe señalar que en estos momentos el INE esta actualizando el marco PIDEH con información del Censo de abril 2002.

1. Censo y Marco de Muestreo de Áreas

Para la creación del marco de muestreo de áreas se utilizó la información recogida del Precenso de Población de 1991. Para ello fue necesario realizar un ajuste geográfico de entidades y actualizar la información cartográfica, como también la condición de las viviendas en lo relativo a las viviendas particulares ocupadas, hogares y personas que las componen.

Las ventajas de la utilización del Precenso en el diseño del marco de muestreo fueron las siguientes:

a) Permitió elaborar en forma más oportuna la confección del nuevo marco para el Programa Integrado de Encuestas de Hogares (PIDEH[1]) por la cercanía al Censo de Población y Vivienda de 1992.

b) Fue posible dirigir a un menor número de empadronadores en relación a lo que requiere el Censo. Ello permitió mejorar la capacitación y lograr un trabajo de terreno más controlado y de mejor calidad en la medición de algunas características de la población.

c) En el formulario correspondiente al área urbana se agregaron variables que permitirían perfeccionar la clasificación socioeconómica de las manzanas, mejorando de este modo la estratificación socioeconómica de los hogares.

d) Finalmente, de la hoja del Precenso se extrajo información a nivel de manzana o entidad, según sea el área geográfica Urbana o Rural, para confeccionar el Marco PIDEH.

Otro de los aspectos de interés para la definición del marco fue la etapa de actualización cartográfica y ajuste geográfico. Esta consistió en aplicar a la cartografía del Precenso todas las modificaciones registradas en los planos censales de actualización antes del levantamiento censal, es decir, aparición o desaparición de manzanas, entidades, etc.

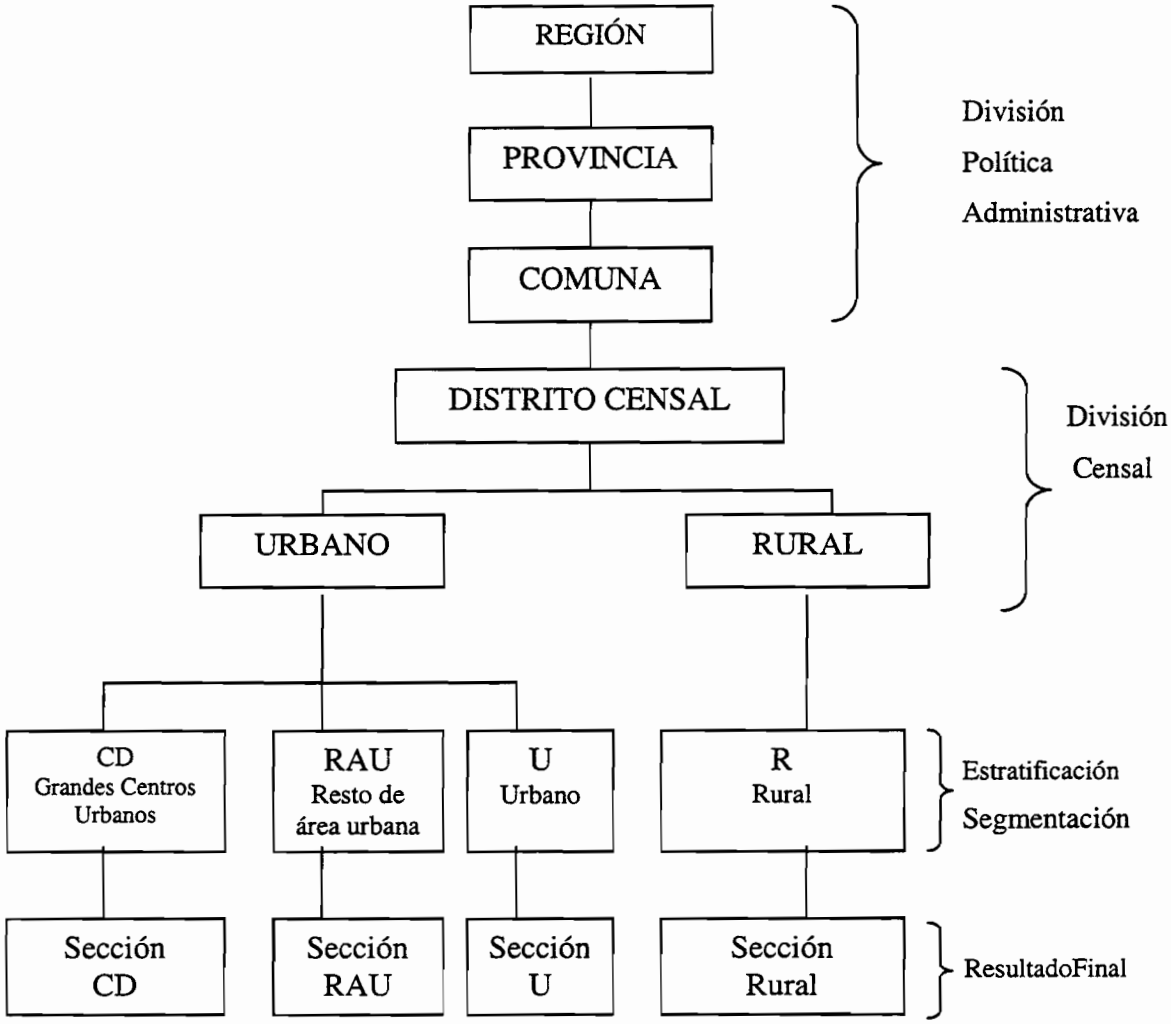
[1] En su última versión (1996) estuvo conformado entre otros por los estudios: Encuesta Nacional del empleo (ENE), Encuesta de Presupuestos familiares (EPF), Estudio sobre el consumo de Drogas (CONACE), Estudio sobre el Comportamiento Sexual (CONASIDA), Calidad de Vida, Estudio sobre el Trabajo a domicilio, etc.

La etapa de ajuste geográfico permitió revisar la cobertura censal verificando y redefiniendo los límites del área urbana y del área rural. Se determinó que algunas entidades (pueblos y aldeas) habían presentado variaciones en la concentración de población y actividad económica que presentaban según los datos del último Censo y debían por lo tanto variar su condición.

Finalmente, se verificaron de acuerdo con los datos recogidos del Censo, aquellas áreas que presentaban viviendas desocupadas (por estar en construcción, en venta o en demolición, etc.) y que posteriormente se encontraron ocupadas.

Por otra parte, el rol que juegan los mapas, planos y fotografías aéreas en un estudio censal, es asegurar el cubrimiento completo del área, la cuenta total de las unidades de investigación dentro de las áreas y mejorar la rapidez y exactitud de las operaciones de campo. El uso de mapas y fotografías aéreas en un marco de muestreo tiene un papel similar, con la diferencia, de que sólo se necesitan mapas de escala pequeña que muestren las Unidades Primarias de Muestreo (UPM). En las etapas siguientes, sólo se necesitarán mapas o fotografías aéreas de escala más grande para las áreas que se hayan escogido en la etapa precedente.

En la siguiente figura, se ilustra un esquema de la estratificación y segmentación, que tiene el marco para llegar hasta las áreas más pequeñas para el empadronamiento.



2. Construcción del Marco de Muestreo de Áreas

A continuación se señalan los puntos que se consideraron para la creación del marco de muestreo de áreas para el PIDEH

2.1 Formación de áreas

Para la elaboración de un marco de muestreo de áreas se requiere de la cartografía censal, complementada con información censal, como el número de viviendas por manzanas, transcritos a los mapas.

Las áreas constituyen una partición del territorio, es decir, no puede haber traslape entre dos o más áreas y la unión del conjunto de áreas debe coincidir con un estrato. En la formación de las áreas se deben conjugar los siguientes elementos:

a) Tamaño del área.

Las áreas deben tener un tamaño homogéneo medido en términos del número de vivienda. La superficie territorial no constituye un criterio para la formación de las áreas.

b) Limite del área

Las áreas deben estar bien definidas y son inamovibles hasta una futura actualización del marco de muestreo. Para estos efectos, sus limites deben ser estables en el tiempo y de fácil identificación en terreno, de manera que el actualizador y el encuestador puedan ubicar sin dificultad el área que ha sido seleccionada y puedan determinar con exactitud que una vivienda se encuentra dentro o fuera de ella.

En las zonas urbanas los limites generalmente corresponden a los ejes centrales de calles. En las rurales en cambio, los limites pueden estar constituidos por lo siguientes elementos:

ríos, quebradas, esteros, cumbres de cerros, barrancos, caminos, canales, líneas de altas tensión, líneas férreas, linderos de predios, prolongación de calles y caminos, paralelas de una camino, etc.

c) Facilidad de acceso a las viviendas de la sección

Al conformar las áreas, en particular en las zonas rurales, se debe analizar la red vial para asegurar la facilidad de acceso a todas las viviendas de cada una de las áreas, a objeto de minimizar los costos de desplazamiento en los procesos de actualización de las áreas y de levantamiento de la encuesta.

2.2 Conformación de estratos del marco de muestreo

Los estratos se definieron según su condición geográfica (División Política Administrativa) y según el numero de vivienda y población que contenían, tomando la información proporcionada por el censo de 1992.

Por consiguiente las definiciones fueron las siguientes:

- Ciudades o Grandes Centros Urbanos: (CD) constituidas por ciudades o conjuntos de ciudades adyacentes con 40.000 o más habitantes
- Resto de Área Urbana: (RAU) formado por conjunto de centros urbanos con menos de 40.000 habitantes.
- Área Rural: (R) integrado por el conjunto de entidades clasificadas como rurales, según si contiene menos de 1.000 habitantes, o entre 1.001 y 2.000 habitantes, con predominio de Población Económicamente Activa (PEA) dedicada a las actividades primarias.

Por lo tanto, el País fue dividido en 146 estratos (Ver anexo 1).

2.3 Unidades de primera etapa (UPE) o secciones

Dadas las características antes mencionadas del marco de áreas, la construcción del marco de muestreo se baso en el material cartográfico nacional, compuesto por los mapas regionales, provinciales, comunales y distritales conteniendo la delimitación de las áreas urbanas y rurales, y la demarcación de las áreas de difícil acceso (ADA)[2].

Con estos antecedentes se procedió a particionar el territorio en distintos estratos y en áreas geográficas o conglomerados denominados Secciones.

Tales secciones se conformaron considerando la cantidad de población y el número de viviendas. El tamaño de cada sección varia entre los distintos estratos que conforman las provincias o regiones del País.

En la construcción de las secciones se consideró además su estabilidad en el tiempo, su fácil identificación en terreno (respetando los limites comunales tanto en las secciones urbanas como rurales), su homogeneidad y su compacidad.

Cabe señalar, que en general la formación de secciones tiende a incrementar la varianza muestral. Sin embargo, su utilización en el diseño de la muestra permite flexibilidad en el número de etapas de selección y reduce sustancialmente los costos asociados a la construcción del marco y la realización del trabajo de campo.

2.4 Tamaño óptimo de las unidades de primera etapa o secciones

Para determinar el tamaño óptimo de las secciones, se tomó en consideración la información recogida de la Encuesta Nacional del Empleo, mediante un estudio del comportamiento de las variables "ocupados" y "desocupados" que medía el nivel de homogeneidad interna y el coeficiente de variación correspondiente al tamaño medio asignado a la Unidad de Primera Etapa (sección) del diseño muestral.

Entonces la determinación del coeficiente de homogeneidad (Rho) se estimó a partir de cada estrato, con la información relativa a:

- El tamaño medio de la sección por estrato, medido en relación al número de viviendas.
- La estimación de ocupados y desocupados por sección y
- Las varianzas inter, intra y total por estrato.

Este coeficiente permite relacionar el error de muestreo y el tamaño de las unidades de primera etapa (UPE) manteniendo constante las demás variables.

2.5 Actualización de las áreas seleccionadas

Las áreas seleccionadas permanecen en la muestra hasta una futura actualización del marco de muestreo. Cabe esperar, que con el transcurso del tiempo el número de viviendas contenidas en cada una de las áreas cambia.

A objeto de identificar y registrar estos cambios las áreas seleccionadas deben ser sometidas periódicamente a un proceso de actualización. Producto de este proceso se dispone de un registro actualizado de viviendas que en ese momento existen en el área, el cual constituye el marco de muestreo para la selección de un nuevo conjunto de viviendas.

Este sistema de actualización permite incorporar periódicamente al marco de muestreo las nuevas viviendas que aparecen en el proceso de expansión urbana y asignarle una probabilidad no nula de pertenecer a la muestra.

---

[2]Areas de difícil acceso, estas áreas son aquellas que por razones de orden climático, topográfico o por ausencia de vías y medios de comunicación expeditos permanecen aislada parcial o totalmente en el año. En ellas habitan aproximadamente 0.8% de la población

La actualización no se realiza simultáneamente en todas las áreas seleccionadas debido básicamente a dos razones:

a) La rotación de viviendas en todas las áreas simultáneamente podría tener un impacto indeseable en las estimaciones de encuestas continuas.

b) El uso eficiente de los recursos aconseja asignar una carga constante y continua al equipo de actualizadores de áreas.

En atención a lo anterior, en los estratos de mayor crecimiento de viviendas y población, cada trimestre se actualiza 1/6 de las áreas, completando así un ciclo de actualización cada seis trimestres. En cambio, en los estratos de menor crecimiento es suficiente actualizar sólo 1/12 de las áreas trimestralmente.

2.6 Selección de una muestra en etapas

La selección de las unidades puede efectuarse de modo que en cada etapa, todas las áreas del universo tengan la misma probabilidad de ser incluida en la muestra o bien con probabilidades proporcionales a algunas medidas de tamaño de la unidad (ejemplo: N° de vivienda, población, etc).

La unidad de primera etapa es proporcional al tamaño y de selección aleatoria y las unidades de segunda etapa con probabilidades iguales y de selección sistemática.

3. Descripción del Diseño Muestral Utilizado para la Encuesta Nacional del Empleo

En la confección del diseño muestral se debe tomar en consideración las variables que se desean investigar. En el caso particular de la Encuesta Nacional del Empleo, ENE, estas se refieren a la evolución de la Fuerza de Trabajo o Población Económicamente Activa (PEA) a nivel nacional y para los distintos niveles de estimación.

El diseño es probabilístico, bi-etápico, estratificado, con probabilidad de selección proporcional al tamaño de la sección en primera etapa y de selección aleatoria, con probabilidad igual en las unidades de segunda etapa o viviendas y de selección sistemática. Los estimadores asociados al diseño son no-autoponderados y se ajustan por una proyección exógena de población calculada con métodos demográficos.

3.1 Población objetivo

La población objetivo de la muestra para la ENE está constituida por toda la población del país residente en viviendas particulares ocupadas. Esta definición excluye a la población que habita en viviendas colectivas, como hospitales, cárceles, conventos, cuarteles y otros, pero incluye a las personas que residen en viviendas particulares dentro de dichos centros, como son los porteros, conserjes y otros. Como así también, en la cobertura geográfica, se excluye la población residente en las áreas de difícil acceso denominadas ADA. Por otra parte, la población objetivo es la que habita todo el territorio continental de Chile.

3.2 Cobertura geográfica y áreas de estimación

Las áreas de estimación cubiertas por la muestra corresponden, en términos generales, a nivel Nacional: Total, Urbano, Grandes Centros Urbanos, Rural; a nivel Regional: Total, Urbano, Rural y Grandes Ciudades, desagregación que no se considera en todas las regiones; y a nivel provincial: Total y Grandes Ciudades.

La necesidad de entregar mayor confiabilidad y precisión en las estimaciones obligó a disminuir las áreas de estimación y aumentar el número de estratos con respecto al marco anterior (PIDEH 1985, Censo 1982). Para ello, se eliminaron las áreas de estimación de los estratos RAU y Rural a nivel provincial y se aumentaron las de Grandes Centros Urbanos, quedando la actual muestra configurada por 79 áreas de estimación y 146 estratos

3.3 Unidades de muestreo y unidades de información

Conservando el número de etapas en que se seleccionó la muestra en el diseño anterior y considerando la eficiencia estadística y continuidad del diseño, se define como unidad de primera etapa a la sección y como unidad de segunda etapa, las viviendas particulares ocupadas que existen dentro de las secciones, siendo las unidades de información o análisis las personas que componen la vivienda, encuestando a todas aquellas que tienen 15 años o más.

El tamaño medio de las secciones por estrato para este marco de áreas es,

Estrato	Pideh - 92	
	$\overline{M}$	$\overline{m}$
CIUDADES	100 -200	8
RAU	60 - 120	12
RURAL	50 - 100	15

3.4 Determinación del tamaño de la muestra

Debido al uso de un diseño de muestra bi-etápico la determinación del tamaño de la muestra se convierte en un trabajo complejo, ya que es necesario incorporar un análisis de precisión de las variables en estudio, tanto en las unidades de primera como de segunda etapa. En este caso, se suma la dificultad asociada al hecho de que la infraestructura muestral es de propósitos múltiples y, por lo tanto, eso hace necesario considerar la posibilidad de que otras variables estén bien medidas.

En este contexto, la determinación del tamaño de la muestra, trata, en primer lugar, de ser compatible con los niveles de estimación y la importancia de los estratos definidos, y con las principales variables que son objeto de estimación.

Para la afijación de la muestra en los distintos estratos se tuvo en cuenta además de la estructura piramidal de los errores de muestreo, otros factores propios de cada estrato tales como:

- El valor de la variable " tasa de desocupación "
- Coeficiente de variación
- Costos de actualización y encuestaje
- La varianza poblacional con el fin de definir y analizar los márgenes o intervalos de error para los distintos estratos y así poder optar por un tamaño adecuado de muestra.

En una primera aproximación se trató de fijar un tamaño de muestra, en términos del total de viviendas, que resultase compatible con la obtención de un error de muestreo de baja magnitud en la estimación de la tasa de desocupación. El tamaño se fijó suponiendo un Muestreo Aleatorio Simple (MAS) de viviendas, el cual se ajustó por un factor que da cuenta del "efecto de diseño" (Deff) en el tamaño, al muestrear por conglomerados (secciones) en lugar de viviendas.

Otro de los antecedentes necesarios de considerar en esta determinación es la cobertura de actualización, con el propósito de identificar viviendas nuevas o desaparecidas.

De esta forma el tamaño muestral Total País logrado según las unidades de muestreo fue de:

Unidades Muestrales	Total	Urbano	Rural
Secciones (UPE)	3477	2934	543
Viviendas (USE)	34269	26124	8145

Para fijar el número de unidades de muestreo de segunda etapa a investigar dentro de cada unidad de primera etapa, se adoptaron varios enfoques, los cuales se basaron, principalmente, en hipótesis acerca de variabilidad entre las dos etapas y en criterios operativos de realización del trabajo en terreno y recolección de datos.

Se consideró la presencia e importancia de los estratos CD a nivel provincial y regional, por cuanto mientras mayor es esa presencia y relevancia, menor es la exigencia de tamaño muestral para los restantes estratos, suponiendo una determinada magnitud de error deseada a nivel regional y/o provincial.

Como normalmente ocurre, otro criterio relevante para determinación del tamaño de muestra fueron los recursos disponibles por el INE (monetarios, materiales y de personal).

Otro de los beneficios de aumentar el número de UPE en la muestra a 3477, estuvo en obtener una mayor dispersión de las unidades y una cobertura mayor de actualización rotando más unidades en cada turno, tema que será tratado más adelante.

3.5 Selección de la muestra

La selección de la muestra se realiza en dos etapas. En la primera etapa se seleccionan las secciones (UPE) pertenecientes a un estrato o área geográfica bien delimitada y cuyo tamaño se expresa en términos de una magnitud de población y un determinado número de viviendas. Y en la segunda etapa se seleccionan viviendas particulares ocupadas existentes dentro de las secciones.

La selección de las unidades de primera etapa se efectuó en forma aleatoria y con una probabilidad de selección proporcional a una medida de tamaño, en particular, el número de viviendas que tenían las mismas al Censo 1992.

Una vez seleccionadas estas unidades, se procedió a actualizar su tamaño -número de viviendas- obteniéndose viviendas actualizadas por sección, quedando así definido el marco muestral para proceder a la selección de las unidades de segunda etapa.

Por su parte, la selección de las unidades de segunda etapa (viviendas dentro de cada estrato) se hizo en forma sistemática con probabilidad igual para las viviendas de cada sección.

El proceso de selección contempla que las unidades de primera etapa, UPE, formen parte de la muestra durante todo el tiempo que el marco esté vigente; mientras que las viviendas que se seleccionan, USE, están sujetas a períodos de rotación a través del tiempo.

3.6 Distribución de la muestra

La muestra total (3477 secciones) se distribuye en un período trimestral, dividiéndose en tres muestras de un tamaño aproximadamente similar, cada una de las cuales se asigna a uno de los tres meses. De esta manera, las consultas de cada una de estas muestras más pequeñas duran un mes.

Así, resulta que todas las viviendas encuestadas en el mes "t", vuelven a serlo en el mes "t+3" y cada vivienda de la muestra es encuestada sólo una vez por trimestre.

Las estimaciones para el período (mes) "t" se realizan utilizando la información de los meses "t", "t-1", "t-2".

3.7 Actualización y enumeración de secciones para rotación de viviendas seleccionadas

La rotación de secciones es un procedimiento que tiene como finalidad mantener actualizada la muestra y evitar el cansancio de los informantes.

Lo anterior consiste en actualizar la enumeración de la sección luego de un cierto período de tiempo, para registrar cambios y posibles crecimientos o decrecimientos que se hayan producido, para posteriormente efectuar una nueva selección de viviendas, con el objeto de incorporar los cambios en la

estimación de las cifras que proporciona la muestra y al mismo tiempo, evitar el "cansancio" de los informantes.

Para efectuar esta rotación, las secciones de la muestra se dividen en "Turnos de Rotación" (TR), los que se estructuran considerando 1/18 del total de secciones urbanas (CD Y RAU) y 1/36 de las secciones rurales de cada región mensualmente. Esto permite rotar la totalidad de la muestra en un período de 18 meses en el caso de las secciones urbanas y en 36 meses para las secciones rurales, con lo que se logra el objetivo de mantenerla permanentemente actualizada y vigente por la incorporación de los cambios que la afectan.

En el período Marzo-Diciembre de 1995, se realizó en todo el país la enumeración base de las secciones seleccionadas en la muestra actualmente en uso, cuyo marco muestral es el Censo de 1992. Actualmente se dispone de archivos computacionales que contienen la enumeración base de todas las secciones seleccionadas en la muestra.

Por lo anterior, el procedimiento de trabajo para actualizar y enumerar las secciones de los turnos de rotación de la muestra, es diferente de la modalidad utilizada en la muestra antigua, en cuanto a las secciones CD y RAU. Las secciones rurales se continuarán trabajando en la forma tradicional. Como se describe a continuación:

a) Enumeración CD o RAU

La enumeración de una sección perteneciente a estratos CD o RAU, bajo el nuevo procedimiento de trabajo, consiste básicamente en verificar lo existente en la enumeración base, con respecto a lo existente en la actualidad, apoyándose en el listado de enumeración base de la sección.

Los cambios producidos se registran por medio de cualquiera de las siguientes acciones: MODIFICACIÓN, COMPLEMENTACIÓN, ELIMINACIÓN, e INSERCIÓN.

MODIFICACIÓN: Consiste en "cambiar", en el mismo formulario de la enumeración base, datos erróneos o que han experimentado variación por el tiempo transcurrido, que son parte de un registro. Nunca se efectúan modificaciones sobre registros completos.

COMPLEMENTACIÓN: La complementación consiste en "agregar", en el listado computacional, información que es parte de un registro y que no aparece impresa debido a que en cada turno de rotación debe ser actualizada.

ELIMINACIÓN: La eliminación consiste en "sacar", del listado de la enumeración base, registros completos porque los edificios, viviendas u hogares que los originaron ya no existen.

INSERCIÓN: La inserción consiste en "incorporar" al listado computacional, registros completos de: edificios, viviendas u hogares que no existían cuando se realizó la enumeración anterior.

b) Enumeración Rural.

La actualización y enumeración de secciones rurales se debe realizar en forma completa, enumerando y actualizando las viviendas. El trabajo se ejecuta de acuerdo al procedimiento utilizado en la muestra antigua. Sin embargo, igualmente se dispone del listado computacional de la enumeración y del croquis con las viviendas punteadas y numeradas, en el período base, con el objeto que sirva únicamente como referencia para lograr un efectivo mejoramiento de la calidad del trabajo original.

4. Avance de la Actualización del Marco de Muestreo de Áreas con Precenso 2001.

En el presente año se esta llevando a cabo la actualización del marco PIDEH´96 con la información del Precenso del 2001 y ajustada con el Censo de población y vivienda del año 2002.

La metodología considera cambios en las áreas Urbanas y Rurales partiendo por la medición de los cambios producidos en los actuales estratos con relación al limite Urbano-Rural definido por el precenso 2001, definiendo a su vez que la Comuna será el área de estimación mínima al que se podrá acceder.

Para tal efecto el nuevo marco propone el distrito en lugar de la comuna como unidad mínima para la definición de sección (áreas) en las zonas Urbanas. Con la opción anterior se podrán definir estratos más pequeños, lo que permitirá capturar más fácilmente el crecimiento habitacional postcensal con la muestra de empleo ENE y con cualquier otra muestra de hogares seleccionadas desde el marco.

Otra de las innovaciones de este marco, es la cartográfica digital para las áreas Urbanas que corresponde al 80% del total País.

Referencias

1. Metodología de la Encuesta Nacional del Empleo 1996, INE-Chile
2. V. Correa, G. Hintze, T. Varela, Junio 1995. El Censo de Población y Vivienda y los Marcos de Áreas. Revista "Estadística y Economía" INE-Chile, capítulo 10, Paginas:89 a 100.
3. Marcos de muestreo y diseños muestrales para Programas Integrados de Encuestas de Hogares. 1987 Naciones Unidas. Programa para el desarrollar la capacidad nacional de efectuar encuestas de hogares.

ANEXO N°1: AREA DE ESTIMACIÓN Y ESTRATOS
PROGRAMA INTEGRADO DE ENCUESTAS DE HOGARES

AREA EST/ESTRATO
TOTAL PAIS
URBANO
GRANDES CENTROS URBANOS
RESTO AREA URBANA
RURAL
I REGION
I URBANO
I CD ARICA
I CD IQUIQUE
I RAU I
I RURAL
II REGION
II URBANO
II CD ANTOFAGASTA
II CALAMA/CHUQUI
II CD CALAMA
II CD CHUQUICAMATA
II RAU II
II RURAL
III REGION
III URBANO
III U PR CHAÑARAL
III CD COPIAPO
III RAU PR COPIAPO
III CD VALLENAR
III RAU PR HUASCO
III RURAL
IV REGION
IV URBANO
IV RURAL
IV PR ELQUI
IV U PR ELQUI
IV COQ/LA SER
IV CD COQUIMBO
IV CD LA SERENA
IV RAU PR ELQUI
IV R PR ELQUI
IV PR LIMARI
IV U PR LIMARI
IV CD OVALLE
IV RAU PR LIMARI
IV R PR LIMARI
IV PR CHOAPA
IV U PR CHOAPA
IV R PR CHOAPA

AREA EST/ESTRATO
V REGION
V URBANO
V RURAL
V PR PETORCA
V U PR PETORCA
V R PR PETORCA
V PR LOS ANDES
V U PR LOS ANDES
V R PR LOS ANDES
V PR SAN FELIPE
V U PR SAN FELIP
V R PR SAN FELIP
V PR QUILLOTA
V U PR QUILLOTA
V CD QUILL-CAL-CR
V RAU PR QUILL
V R PR QUILLOTA
V PR VALPARAISO
V U PR VALPO
V GR VALP
V CD VALPARAISO
V CD VIÑA DEL MAR
V CD QUILPUE
V CD VILLA ALEMANA
V RAU PR VALPO
V R PR VALPO
V PR SAN ANTONIO
V U PR SAN ANTON
V CD SAN ANTONIO
V RAU PR SAN AN
V R PR SAN ANTON
VI REGION
VI URBANO
VI RURAL
VI PR CACHAPOAL
VI U PR CACHAPOAL
VI CD RANCAGUA
VI RAU PR CACHAP
VI R PR CACHAPOAL
VI PR COLCHAGUA
VI U PR COLCHAGUA
VI CD SAN FERNANDO
VI RAU PR COLCHAGUA
VI R PR COLCHAGUA
VI PR CARD CARO
VI U PR CARD CARO
VI R PR CARD CARO
VII REGION
VII URBANO
VII RURAL
VII PR CURICO
VII U PR CURICO
VII CD CURICO
VII RAU PR CURICO
VII R PR CURICO
VII PR TALCA
VII U PR TALCA
VII CD TALCA
VII RAU PR TALCA
VII R PR TALCA
VII PR LINARES
VII U PR LINARES
VII CD LINARES
VII RAU PR LINAR
VII R PR LINARES
VII PR CAUQUENES
VII U PR CAUQUENES
VII R PR CAUQUENES

AREA EST/ESTRATO
VIII REGION
VIII URBANO
VIII RURAL
VIII PR ÑUBLE
VIII U PR ÑUBLE
VIII CD CHILLAN
VIII RAU PR ÑUBLE
VIII R PR ÑUBLE
VIII PR CONCEPCION
VIII U PR CONCEP
VIII GR CONCEP
VIII CD CONCEPCION
VIII CD TALCAHUANO
VIII CD PENCO
VIII LOTA/COR
VIII CD LOTA
VIII CD CORONEL
VIII RAU PR CONCEP
VIII R PR CONCEP
VIII PR ARAUCO
VIII U PR ARAUCO
VIII R PR ARAUCO
VIII PR BIO-BIO
VIII U PR BIO-BIO
VIII LOS ANGELES
VIII RAU PR BIO-B
VIII R PR BIO-BIO
IX REGION
IX URBANO
IX RURAL
IX PR MALLECO
IX U PR MALLECO
IX CD ANGOL
IX RAU PR MALLECO
IX R PR MALLECO
IX PR CAUTIN
IX U PR CAUTIN
IX CD TEMUCO
IX RAU PR CAUTIN
IX R PR CAUTIN
X REGION
X URBANO
X RURAL
X PR VALDIVIA
X U PR VALDIVIA
X CD VALDIVIA
X RAU PR VALDIV
X R PR VALDIVIA
X PR OSORNO
X U PR OSORNO
X CD OSORNO
X RAU PR OSORNO
X R PR OSORNO
X PR LLANQUIHUE
X U PR LLANQUIH
X CD PUERTO MONTT
X RAU PR LLANQ
X R PR LLANQUIH
X PR CHILOE/PAL
X U PR CHILOE/PAL
X R PR CHILOE/PAL
XI REGION
XI URBANO
XI CD COIHAIQUE
XI CD PTO AISEN
XI RAU XI
XI RURAL
XII REGION
XII URBANO
XII CD PUNTA ARENAS
XII RAU XII
XII RURAL

AREA EST/ESTRATO
RM REGION
RM URBANO
RM RURAL
RM PR SANTIAGO
RM SANTIAGO
RM INDEPENDENCIA
RM RECOLETA
RM EST CENTRAL
RM CONCHALI
RM HUECHURABA
RM QUILICURA
RM RENCA
RM QUINTA NORMAL
RM PUDAHUEL
RM CERRO NAVIA
RM LO PRADO
RM MAIPU
RM CERRILLOS
RM LA CISTERNA
RM LO ESPEJO
RM EL BOSQUE
RM LA GRANJA
RM LA PINTANA
RM SAN RAMON
RM SAN MIGUEL
RM P. AGUIRRE C.
RM SAN JOAQUIN
RM LA FLORIDA
RM ÑUÑOA
RM MACUL
RM PEÑALOLEN
RM LA REINA
RM PROVIDENCIA
RM LAS CONDES
RM VITACURA
RM LO BARNECHEA
RM PR CHACABUCO
RM U PR CHACABUCO
RM R PR CHACABUCO
RM PR CORDILLERA
RM U PR CORDILL
RM CD PUENTE ALTO
RM RAU PR CORDI
RM R PR CORDILL
RM PR MAIPO
RM U PR MAIPO
RM CD SAN BERNARDO
RM RAU PR MAIPO
RM R PR MAIPO
RM PR MELIPILLA
RM U PR MELIPILLA
RM CD MELIPILLA
RM RAU PR MELIPILLA
RM R PR MELIPILLA
RM PR TALAGANTE
RM CD PADRE HURTADO
RM RAU PR TALAGANTE
RM R PR TALAGANTE
RM GRAN SANTIAGO

DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES

DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA (DANE)
COLOMBIA

ÍNDICE

Página

Introducción 145
I.- Antecedentes 145
II.- Ampliación de la Muestra Maestra 145
III.- Sistema de Información Geoestadística 149

Introducción

A continuación se presentan los principales aspectos referentes a la experiencia de Colombia en la utilización de los censos de población como insumo principal en la construcción de los marcos de muestreo para las encuestas a hogares, vale decir, Encuesta Continua de Hogares, Encuesta de Calidad de Vida, Encuesta de Ingresos y Gastos y Encuesta de Trabajo Infantil. También se hace una explicación de la ampliación de la muestra maestra de propósitos múltiples del 2001 así como de todos los aspectos cartográficos relacionados con la construcción de las unidades secundarias de muestreo como parte fundamental en su implementación.

I. Antecedentes

Hasta la fecha se han realizado 16 censos nacionales de población y cinco de vivienda. El último de ellos en el año de 1993 y se tiene programado para el 2003 la realización del XVII Censo Nacional de Población y VI de vivienda. El marco de muestreo utilizado actualmente corresponde al generado precisamente con base en el censo del año 1993. Con este marco se realizaron las encuestas de Calidad de Vida de 1997 y las encuestas de hogares para la medición de la fuerza laboral, así como la Encuesta de Trabajo Infantil para el año 2001 y 2002 y el Observatorio de Seguridad para la ciudad de Barranquilla.

El mayor uso es el dado a la Encuesta Continua de Hogares, donde se mide la fuerza laboral. El sistemático perfeccionamiento de su metodología y esquemas operativos ha sido una constante en los últimos años. En 1996 – 1997 se realizaron importantes innovaciones: se aumentó el tamaño de la muestra para incrementar la precisión de las estimaciones; se ajustó la distribución departamental, la de las grandes ciudades y la regional, para garantizar un error de muestreo similar para estos niveles de desagregación, en los indicadores mas importantes, se incorporaron los materiales de la muestra diseñada para la evaluación de cobertura del censo 93, que por su diseño, tamaño y estructura podía ser utilizada con el carácter de muestra maestra.

En 1998, en atención a las recomendaciones del señor Ralph Hussmans, Experto en Estadísticas Laborales de la OIT, se inició el proceso de conversión de la Encuesta Nacional de Hogares (ENH) de una investigación "transversal", con medición trimestral de los indicadores de empleo, a una de tipo "continuo", para una captación del comportamiento integral de los fenómenos laborales, durante todo el año. La primera etapa de conversión, llevada a cabo exitosamente, comenzó en la primera semana de 1999 con un esquema de semanalización de la muestra y de la encuesta trimestral de Hogares de Bogotá, D.C., paralelo al de la encuesta transversal tradicional al cual se incorporaron en el segundo semestre Medellín, Cali, Barranquilla y Bucaramanga.

Desde la primera semana del año 2000 se agregaron a la fase experimental continua las restantes grandes ciudades y áreas metropolitanas de la Encuesta transversal trimestral y se continuó, al mismo tiempo, con la encuesta transversal. La simultaneidad de los esquemas continuo y trasversal, se mantuvo durante todo este año.

En enero del 2001 se dio inicio a la encuesta continua para el resto urbano y la zona rural con submuestras replicadas mensuales del diseño construido a partir de la muestra maestra de este resto de país. Finalmente a partir de este año entró en plena vigencia y de manera oficial la Encuesta Continua de Hogares.

II. Ampliación de la Muestra Maestra

Debido a que cada vez se hace mas evidente el requerimiento de los usuarios de la información de las encuestas de hogares por muestreo, de estimaciones con el máximo de desagregación geográfica posible, se requirió del diseño de una gran muestra que tuviese gran flexibilidad en la planeación del tamaño y características de las muestras específicamente requeridas por las diversas investigaciones nacionales del país, así como la posibilidad de acumulación de resultados de varios estudios que tengan componentes o

módulos similares y que hayan sido realizados simultáneamente. Es así como la muestra diseñada para la realización de la Encuesta Postcensal de Cobertura del Censo de Población y Vivienda del 93, se diseñó con el carácter de Maestra y se implementó para la Encuesta Nacional de Hogares . En el 2001 se realizó una ampliación de esta muestra maestra bajo las siguientes consideraciones:

- **Definición integral del tamaño y composición de la muestra que sería utilizada en los próximos años**, con el fin de permitir la planeación de materiales y operaciones más eficientes y baratos, de garantizar una mejor calidad y precisión de los resultados, y de permitir la confiable comparación y seguimiento de los indicadores generados a través del tiempo. Los cálculos para la estimación anual de los indicadores laborales a nivel departamental, señalan un requerimiento aproximado de 9,000 segmentos; un esquema de rotación a tres años, eleva esta cifra a alrededor de 27,000 segmentos. Los estudios coyunturales podrían exigir entre 500 y 2,5000 segmentos cada uno.

- **Ampliación del número de estratos y unidades primarias de muestreo (UPM)** para poder completar 12 submuestras nacionales replicadas (una por mes), y para facilitar la eventual obtención de indicadores departamentales. Se completaron 205 UPM (con 244 mpios.), frente a 165 existentes, (con 166 mpios.), con un total de 228 estratos (algunas UPM grandes tienen 2 o 3 estratos). Consiguientemente, se reestructuró toda la actual estratificación, con base en la última información disponible sobre las variables de clasificación (Censo 93 y sus proyecciones). **Las 13 grandes ciudades con sus áreas metropolitanas, no están contempladas en esta reestructuración, pues son autorrepresentadas .**

  El número 228 es modular no sólo para la conformación de las 12 submuestras replicadas mensuales de 19 estratos cada una, sino que permite la conformación de otro conjunto de clasificaciones o submuestras nacionales replicadas para los estudios coyunturales: 4 submuestras de 57 UPM cada una y 3 submuestras de 76 UPM cada una. Estas clasificaciones permiten, a su vez, ofrecer un variado menú de submuestras nacionales de la siguiente magnitud, según el número de hogares requeridos:

Requerimiento muestral de un estudio en particular		N° de Estratos	N° de Submuestras combinadas
Segmentos (10 a 20 por estr.)	Hogares		
570 – 1,140	5,700 – 11,400	57	1
760 – 1,520	7,600 - 15,200	76	1
1140 – 2,280	11,400 – 22,800	114	2 (de 57 UPM)
1520 – 3,040	15,200 – 30,400	152	2 (de 76 UPM)
1,710 – 3,420	17,100 – 34,200	171	3 (de 57 UPM)

  La decisión del número de segmentos por estrato, dentro del rango, dependerá de consideraciones operativas, de costo y sobre el efecto esperado de los conglomerados (Deff).

- **Mantenimiento de buena parte de las UPM de la muestra actual**, para poder seguir utilizando los materiales muestrales disponibles, y el pertinente ajuste de las probabilidades de selección y de los factores básicos de expansión. De las 205 UPM de la nueva muestra, 129 (63%) son de la muestra vigente (Cuadro 1).

Actividades

Las actividades realizadas para la ampliación de la muestra comprenden:

1. Redefinición de las Unidades Primarias de Muestreo (UPM). Se estableció un tamaño poblacional total mínimo de 3,000 habitantes, y de 2000 habitantes en la cabecera municipal, para poder generar

los segmentos muestrales requeridos por las encuestas del próximo futuro. Integración con UPM vecinas de similares características de las que no cumplían los anteriores requisitos.

2. Revisión de las variables de estratificación con base en la más reciente información disponible, y verificación de que las variables confirmadas tuvieran la menor correlación entre sí.
   Confirmación como variables claves de: Región, departamento, tamaño del estrato, población total de cada UPM, proporción cabecera -resto y proporción con necesidades básicas insatisfechas (NBI).

3. Elaboración de 228 estratos aplicando como herramienta de gran ayuda para su optimización el procedimiento "Clusters" del paquete SAS.

4. Actualización de los parámetros del cálculo del tamaño muestral. Verificación del efecto de los conglomerados en el diseño (deff), con base en los resultados de las etapas 105 y 109 de la Encuesta Nacional de Hogares (ENH). Utilización, como medida de seguridad, del deff más alto observado en cada departamento. Cuantificación del aporte en la varianza total de las estimaciones, de la varianza inter e intra UPM.

5. Realización de cálculos exploratorios de tamaño de muestra a nivel departamental para estimaciones puntuales de los indicadores de interés, particularmente de la tasa de desempleo, con diferentes alternativas de precisión (ESrel = 0.01 a 0.12), y para la estimación de cambios en el tiempo de tales indicadores (cambios de 0.5, 1.0 y 1.5 puntos porcentuales).

6. Adopción de tamaños muestrales guía para la orientación del subsiguiente trabajo, así:

   - 2,400 segmentos anuales (es la cifra actualmente vigente y no incluye grandes ciudades), o sea 200 segmentos mensuales para apoyar la ECH en su versión 2001-2002. Cada segmento tiene en promedio 10 hogares.
   - 9,600 segmentos anuales, que es cuatro veces la cifra anterior, lo cual significa 800 segmentos mensuales, para permitir la estimación del desempleo <u>anual</u>, por departamento, con un ESrel = 0.04 – 0.05
   - 38,400 segmentos totales; 28,800 de ellos para permitir la rotación anual completa de los 9,600 segmentos por un lapso de tres años, y otros 9,600 para facilitar hasta tres estudios coyunturales en el mismo año, cada uno con un requerimiento hasta de 2,400 segmentos.

   Las cifras anteriores son ajustables, por supuesto, ante un raciocinio diferente por parte del equipo directivo del DANE.

7. Clasificación de los 228 estratos en que está dividido el país en 12 "subgrupos replicados" de 19 estratos, es decir, "idénticos" desde el punto de vista de las variables de estratificación. La metodología garantiza óptima representatividad nacional de cada uno de los subgrupos replicados, separadamente, a través de las UPM que los representa. Hablamos entonces de 12 submuestras replicadas, para las encuestas mensuales de la Encuesta Nacional de Hogares.

   Para esta clasificación se aplicó la técnica de "Selección Controlada" (Goodman y Kish, 1950) que, en este caso, optimiza la balanceada composición de las submuestras, vale decir, minimiza la varianza intersubmuestra, y maximiza la varianza intrasubmuestra (lo contrario de la estratificación). Conviene aclarar que esta técnica, diseñada para optimizar el efecto de la estratificación (reducir el error de muestreo, minimizando varianza intra estrato y maximizando varianza inter estrato), se adaptó en este caso para producir el efecto contrario.

8. Distribución de los 2,400 segmentos financiados de la muestra anual actual de la ENH, en 12 paquetes, con la siguiente secuencia:

   - Distribución del total de segmentos (2,400) en los 23 departamentos. Los 81 segmentos rurales de las 13 grandes ciudades se incluyeron en esta distribución. La distribución se hizo en función

del mismo Error Estándar Relativo (coincidió un ESrel = 9 –10%), para fenómenos laborales con una probabilidad de ocurrencia de p = 0.10.

- Distribución por UPM de la muestra maestra de los segmentos de cada departamento, proporcionadamente a la población de los estratos respectivos, y distribución cabecera – resto en cada UPM, conforme a la composición de sus estratos. Con este procedimiento se garantizó probabilidad igual en todos los segmentos y hogares de cada departamento.
- Posterior redondeo en 200 de los segmentos resultantes en cada una de las submuestras nacionales de UPM, proporcionalmente al tamaño de los estratos respectivos. Dichos segmentos resultantes de la distribución del paso anterior fluctuaron ligeramente alrededor de 200 en las 12 submuestras. La igualación en 200 tiende a generar similar error de muestreo en las submuestras para fenómenos de la misma probabilidad de ocurrencia, con las consiguientes ventajas adicionales para la planeación e implementación de los operativos de campo mensuales.

9. Construcción del Marco de Muestreo de las Unidades Secundarias en los municipios seleccionados diferentes a los Autorepresentados: Decisión estructurada sobre los procesos de selección y segmentación de las Unidades Terciarias de Muestreo (UTM), correspondientes a secciones o sectores censales. Para las zonas rurales, el DIG (Dirección de Información Geoestadística del DANE-) hace la segmentación sobre fotos aéreas ampliadas, utilizando tanto lo observado en dichas fotos como la clasificación catastral sobre uso de las edificaciones (vivienda u otro fin); para las zonas urbanas, las regionales del DANE hacen el listado y segmentación de las manzanas seleccionadas por el Equipo de Diseños Muestrales, sobre el material geográfico generado por DIG. La siguiente es la secuencia del trabajo:

- Verificación de que las secciones censales (o sectores) tengan entre 120 y 200 viviendas (160 en promedio), es decir, 4 Medidas de Tamaño (MT) de 40 hogares en promedio, por cada segmento a ser seleccionado, para garantizar el máximo tamaño potencial previsto en la muestra maestra. Si en un sector sólo se va a seleccionar un segmento, o ya existe seleccionado uno, su tamaño debe ser mínimo de 16 MT, si se van a seleccionar 2 MT, debe tener un mínimo de 32 MT, y así sucesivamente. Si no se cumple este requisito, se deben juntar los sectores o secciones insuficientes, con sectores o secciones vecinos.
- En las nuevas UPM, o en los procesos de ampliación de la muestra de las UPM antiguas, seleccionar tantos sectores o secciones (la máxima desagregación existente) como nuevos segmentos se requieran con probabilidad proporcional al número de MT existentes.
- En los sectores o secciones seleccionados, identificación del número de MT existentes (de $\bar{x}$ = 40 hogares). En las zonas urbanas realizar esta identificación a nivel de manzana. Las manzanas que no tengan una MT, por lo menos, se juntarán con una vecina de la misma sección. Aunque se espera que las secciones rurales tengan por lo menos una MT, las que no cumplan con este requisito deben integrarse con una contigua.
- Diseñar una instrucción de submuestreo para los nuevos sectores o secciones seleccionados. Para los sectores urbanos, selección sistemática, con un intervalo constante, de tal manera que el número de selección resultante de la aplicación del intervalo de muestreo, identifique directamente la manzana (o manzanas colapsadas) seleccionadas. Estas áreas últimas muestreadas, cada una de las cuales contiene por lo menos una MT con un promedio de 40 hogares, son las que deben ser visitadas por el personal de las regionales, para su listado y partición en segmentos de 10 hogares en promedio. Finalmente, asignación aleatoria del orden en que serán utilizados esos segmentos.

  Para las zonas rurales, subdivisión de los sectores en tantas partes como medidas de tamaño existentes, submuestreo aleatorio de una de ellas, partición en segmentos de 10 viviendas en promedio, y asignación aleatoria del orden en que van a ser utilizados en las próximas encuestas.

10. Construcción del Marco de Muestreo de las Unidades Secundarias en los municipios Autorepresentados: Las ciudades que se venían trabajando en la Encuesta de Hogares Tradicional, se introdujeron en la muestra como ciudades autorrepresentadas con las cuales se siguen dando estimaciones de las variables de interés de acuerdo a la periodicidad establecida. Para la estratificación y selección de la muestra, las áreas geográficas de estas ciudades con sus áreas metropolitanas se organizaron, según los principios cartográficos establecidos, en sectores, secciones y manzanas, definidos así:

    - Sector: es un área cartográfica delimitada, compuesta por secciones y manzanas; lo conforman uno, dos o más barrios o urbanizaciones, aunque se encuentran casos en que el sector está circunscrito a una parte de un barrio grande.
    - Sección. Es una subdivisión del sector equivalente a 18 manzanas aproximadamente en la parte urbana
    - Manzana. Lote de terreno edificado o sin edificar, delimitado por vías de tránsito vehicular o peatonal de carácter público, es decir por calles, carreras, diagonales, transversales, avenidas o pasos peatonales. Una manzana puede estar delimitada igualmente por un río, una quebrada o un canal siempre y cuando estos elementos sean de carácter permanente y se localicen fácilmente en campo.

Al marco de muestreo, con la información del número de viviendas a nivel de manzana se le adicionó el archivo con la estratificación realizada por cada municipio, con base en la "Metodología de la estratificación de viviendas urbanas en los municipios" con la cual se estratifican las manzanas en 6 estratos socioeconómicos teniendo en cuenta los siguientes factores:

- Características físicas de las viviendas.
- Disponibilidad de servicios públicos domiciliarios.
- Estado de las vías adyacentes y demás elementos de la infraestructura.
- Facilidades recreacionales y zonas verdes.
- Naturaleza y calidad de los medios de transporte.
- Uso y concentración comercial.

Como insumo del XVI Censo Nacional de Población y V de Vivienda, realizado en 1993, se obtuvo el marco de muestreo en el cual esta consolidado el total de personas, hogares y viviendas a nivel de manzana para las cabeceras municipales y a nivel de sector cartográfico para la parte resto del municipio. Las variables de identificación que componen estos archivos llamados VIHOPE (viviendas, hogares y personas) son: Departamento, municipio, clase, sector rural, localidad, sector urbano, sección, manzana, viviendas.

El marco permite tener la información a un nivel de agregación mayor como lo es el municipio con lo cual se trabaja en la primera etapa de selección de la Muestra Maestra

III. Sistema de Información Geoestadística

El marco de muestreo está soportado sobre la cartografía utilizada para la realización del XVI Censo Nacional de Población y V de Vivienda de 1993, donde el DANE consciente de la importancia de contar con una base cartográfica actualizada y aprovechando las posibilidades ofrecidas por la tecnología, llevó a cabo el Proyecto de Actualización y Sistematización de la Cartografía Censal, cuyo objetivo fue actualizar la cartografía censal, para satisfacer los requerimientos censales y estadísticos post-censales de

la Entidad, así como ofrecer la posibilidad de asociar información estadística a la información geográfica, mediante la conformación del DIG[1].

El desarrollo del Proyecto inició con la digitalización de la información básica de la cartografía rural, para lo cual se utilizó como base la cartografía del IGAC[2], escala 1:25.000 en el 51.19% del área total del territorio nacional, 1:50.000 en el 9.13% y 1:200.000 en el 30.53%. Esta captura se realizó por planchas y se capturaron cuatro niveles de información básica: ríos, vías, curvas de nivel y toponimia. El trabajo de actualización en las áreas rurales se realizó mediante la georreferenciación de puntos de referencia[3] y vías, verificación de nombres de ríos y quebradas, verificación de la existencia de centros poblados, actualización de los tipos de vías y ajustes a la sectorización censal, la cual previamente había sido delimitada sobre la cartografía base (trabajo de oficina), garantizando que los sectores censales correspondieran a límites identificables en el terreno.

Para las áreas urbanas se tomó como base la cartografía más actualizada, la cual podía corresponder a planos producidos por las oficinas de Planeación a nivel municipal y/o departamental, por el IGAC y/o DANE. Adicionalmente, se agregaron las novedades encontradas, mediante la realización de levantamientos topográficos con brújula y cinta, la verificación y complementación de nomenclatura vial y de sitios de interés. Para aquellas cabeceras municipales y/o centros poblados que no tenían plano base, se elaboró el plano respectivo, mediante un levantamiento topográfico y la correspondiente georreferenciación[4].

El trabajo inicial de captura de la información básica del área rural y urbana, se realizó durante los años 1992 y 1993, en los años posteriores se realizó el empalme, complementación y estructuración topológica de la cartografía básica complementaria, así como la definición de la estructura para el almacenamiento de la información dentro de una base de datos, la estructuración y adecuación del marco geoestadístico nacional.

Lo anterior le permite al DANE y al país contar hoy con el marco geoestadístico nacional y una información cartográfica de referencia para las áreas urbanas y rurales del territorio nacional, que se encuentran digitalizados y estructurados topológicamente, para permitir su organización dentro de una base de datos gráfica, facilitando los procesos de actualización y mantenimiento de dicha información, a través de operativos de campo, la generación de productos derivados, y la integración con la información estadística que produce la Entidad, dentro del Sistema de Información Geoestadística, tal que facilite el desempeño de las diferentes tareas y procesos tendientes a generar la información estadística estratégica del país.

El Marco Geoestadístico Nacional, es un sistema diseñado para referenciar correctamente la información estadística con los lugares geográficos correspondientes. Esta constituido por la división político administrativa, departamentos, municipios, y sectores geográficos definidos con fines estadísticos, los cuales se encuentran delimitados principalmente por accidentes naturales identificables en terreno. Su actualización y mantenimiento permanentes son de gran relevancia para garantizar la permanencia en el tiempo del Sistema, así como su funcionalidad.

El Marco Geoestadístico, se encuentra complementado con información cartográfica básica para cada municipio, tanto para las áreas urbanas (cabeceras municipales y centros poblados) y rurales. Para las cabeceras municipales y centros poblados la cartografía se encuentra en escala 1:5000 con información hasta nivel manzana, incluyendo sitios de interés y nomenclatura vial. Para las zonas rurales se cuenta con información básica de ríos, vías, toponimia y curvas de nivel, tomada de la base IGAC.

---

[1] DIG Dirección de Información Geoestadística

[2] Instituto Geográfico Agustín Codazzi

[3] Corresponden a sitios de interés locales que facilitan la orientación y ubicación en campo, tales como: Fincas, peajes, escuelas, puestos de salud, estaciones de servicio, instalaciones agroindustriales, iglesias, restaurantes y hoteles entre otros.

[4] Esta georreferenciación se hizo utilizando GPS, Sistema de Posicionamiento Global, con el fin de asignar las correspondientes coordenadas espaciales.

La base de datos gráfica del sistema contiene el marco geoestadístico nacional, es decir, cuenta con los límites de todos los departamentos y municipios del territorio nacional, además de la identificación de todas las cabeceras municipales y de los centros poblados[5]; para éstos últimos se cuenta con información cartográfica hasta nivel manzana. Lo anterior permite realizar consultas que involucran la combinación de datos físicos de la zona y su correspondiente información estadística. Por ejemplo, teniendo la localización de las fuentes para la Encuesta de Comercio y conociendo el número de encuestadores disponibles, se puede hacer la organización de las correspondientes rutas para la entrega y recolección de los formularios. Otro ejemplo, puede ser el de las citaciones para los empadronadores el día del Censo Nacional, a los cuales se les asignan áreas de trabajo próximas a sus lugares de residencia.

[5] Centro poblado, se refiere a los caseríos, inspecciones de policía y corregimientos pertenecientes al área rural del municipio y que están conformados por 20 o más viviendas contiguas o adosadas

DISEÑO Y CONSTRUCCIÓN DE MARCOS MUESTRALES PARA ENCUESTAS DE HOGARES

GISELLE ARGÜELLO VENEGAS

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS (INEC)
COSTA RICA

ÍNDICE

Página

Introducción 155
I.- Marco Muestral de Viviendas 1986 155
II. - Marco Muestral de Viviendas 1998 157
III.- Diseño de la Muestra 158
3.1 Población y cobertura 158
3.2 Dominios de estudios 158
3.3 Unidades de muestreo 158
3.4 Definición de la muestra 158
3.4.1 Determinación del tamaño de muestra 159
3.4.2 Distribución de la muestra 160
3.5 Selección de la muestra 161
3.5.1 Selección de Unidades Primarias de Muestreo (UPM) 161
3.5.2 Selección de Unidades Secundarias de Muestreo (USM) 162
3.6 Rotación de la muestra 162
3.7 Factor de expansión 162
3.8 Ajustes al factor de expansión 163
Ajustes por actualización 163
Ajustes por tasa de respuesta 164
Ajustes por proyecciones de población 165
3.9 Cálculo de errores muestrales 165

Introducción

Costa Rica ha realizado nueve Censos Nacionales de Población y cinco Censos Nacionales de Vivienda. Los más recientes se realizaron en junio del 2000 después de 16 años de la realización de los últimos censos de población y vivienda.

En cuanto a Encuestas de Hogares, Costa Rica cuenta con una larga trayectoria que inicia en 1965 con la realización de la Encuesta Centroamericana de Hogares (ECAH). Hasta 1986 las encuestas de hogares se realizaron de forma discontinua e independientes, en esta fecha se creó el Programa Permanente de Encuestas de Hogares con el fin de mejorar e integrar las encuestas que se habían venido realizando de forma aislada y fortalecer la capacidad técnica de la entonces Dirección General de Estadística y Censos (DGEC) para ejecutar un programa continuo de Encuestas de Hogares. Este programa se ha ejecutado de forma ininterrumpida con una encuesta puntual a julio de cada año e incluye una serie de módulos especiales como son: vivienda, salud, consumo aparente, menor trabajador, tercera edad, telecomunicaciones, etc.

El Marco Muestral de Viviendas (MMV) para la ejecución de este programa se construyó a partir de la información de los Censos Nacionales de Población y Vivienda de 1984. En 1998 a raíz de la cancelación de los Censos Nacionales programados para 1996 y dado que el MMV presentaba graves problemas de desactualización, se consideró urgente la construcción de un nuevo MMV a partir de la actualización cartográfica preparada para los censos, la cual incluía un conteo del número de viviendas por segmento censal.

Con la realización de los Censos del 2000 se iniciaron las actividades preparatorias de un nuevo MMV y del diseño de un Sistema Integrado de Encuestas de Hogares que permita coordinar, complementar, relacionar y sincronizar en el tiempo las encuestas de hogares que de él se deriven.

I. Marco Muestral de Viviendas 1986

El MMV de 1986 se construyó a partir de la base cartográfica y de la información generada por los Censos de Población y Vivienda de 1984. Las áreas de enumeración o segmento censal se utilizaron para la conformación de las unidades de un Marco Muestral de Áreas. Estos segmentos representan áreas geográficas que comprenden en promedio un número de 60 viviendas en la zona urbana y 40 viviendas en la zona rural.

Se realizó una actualización inicial del marco con el fin de incorporar los principales cambios desde la fecha de realización del censo (1984) al momento de la construcción del marco. Posterior a esta actualización se realizó una división y fusión de segmentos con el fin de homogenizar su tamaño y crear unidades del marco los suficientemente grandes para cubrir la rotación de la muestra durante todo el programa de encuestas. Los criterios para dividir o fusionar los segmentos son los siguientes:

	Mínimo (L_1)	**Máximo (L_2)**
Urbanos	Menos de 40 viviendas	Más de 99 viviendas
Rural	Menos de 29 viviendas	Más de 79 viviendas

Del Marco Muestral fueron excluidos los segmentos de deficil acceso y de escasa densidad poblacional, en total 202 segmentos que representaron el 1.52% de la población total.

Las variables incluidas en el marco para su estratifiación fueron variables geográficas: región, provincia, cantón, distrito y grado de urbanización. La asignación del grado de urbanización comprende cuatro categorías: urbano, periferia urbana, rural concentrado y rural disperso y los criterios de asignación fueron los siguientes:

- **Urbano:** corresponde a los centros administrativos de los cantones del país, o sea parte o todo el distrito primero, además de otras áreas adyacentes. Estas áreas fueron demarcadas a priori con criterios físico funcional, tomando en cuenta elementos tangibles como cuadrantes, calles, aceras, luz eléctrica, servicios urbanos, etc.

- **Periferia Urbana:** son aquellas zonas que pertenecen a las áreas localizadas entre el límite del cuadrante urbano y la poligonal involvente del área urbana. La poligonal envolvente consiste de una línea imaginaria que encierra, tanto el cuadrante urbano, como las áreas adyacentes del mismo, que aunque no se encuentran dentro del cuadrante urbano, representan características similares a las de estas zonas. Estas áreas se consideran como de transición de rural a urbano.
- **Rural Concentrado**: son aquellos centros poblados no ubicados en el área urbana en las que el suelo está ocupado predominantemente por actividades no agropecuarias; tienen de 50 o más viviendas agrupadas o contiguas con distancias de no más de 20 metros entre sí; disponen de servicios de infraestructura como electricidad domiciliaria, agua potable, teléfono, escuelas, iglesia, centros de salud, etc.
- **Rural Disperso**: comprende las áreas no incluidas en las categorías anteriores.

Para efectos de dicotomizar esta variable el grado de urbanización urbano y la periferia urbana se unen como zona urbana y el rural concentrado y el rural disperso como la zona rural.

Además de estas variables se incluyó un indicador del nivel socioeconómico. Este indicador se construyó a partir de un método indirecto de estimación del ingreso, se utilizó un modelo de regresión lineal múltiple cuyas variables independientes provienen de los Censos Nacionales de Población y Vivienda y la Encuesta de Hogares de 1984. La información cuantitativa de la estimación del ingreso por vivienda (hogar principal), sirvió de base para estratificar los segmentos del área urbana en el marco. El límite de cada estrato se determinó aplicando el método de Dalenius. Como resutado de esta estratificación, cada segmento del área urbana quedó clasificado en alguno de los siguientes niveles socioeconómicos:

- **Bajo:** segmentos con hogares de ingreso menor a 14 000 colones.
- **Medio:** segmentos con hogares de ingreso entre 14 000 y menos de 29 000 colones.
- **Alto:** segmentos con hogares de ingreso mayor a 29 000 colones.

En un inicio se consideró dos tipos de actualizaciones del Marco Muestral.: una de carácter restringido únicamente para los segmentos de la muestra y una de carácter general a todos los segmentos del marco por medio de un marco de nuevas construcciones. Esta última actualización no fue posible llevarla a la práctica. En cuanto a la actualización de los segmentos de la encuesta, el supervisor de trabajo de campo asignaba una calificación al segmento de acuerdo a los cambios observados:

0 Sin cambios

1 Sin cambios significativos (1 o 2 viviendas)

2 Cambios pequeños (menos del 10% de viviendas)

3 Cambios moderados (del 10% al 49% de viviendas)

4 Cambios muy grandes (más del 50% de viviendas)

5 Cambios cartográficos.

Si el segmento se calificaba con código 3, 4 o 5 se efectuaba un levantamiento completo del segmento. Después de concluir la construcción del Marco Muestral de Áreas, la siguiente etapa fue la aplicación del diseño muestral respectivo para las Encuestas de Hogares de Propósitos Múltiples (EHPM), la Encuesta Nacional de Ingresos y Gastos y la Encuesta Continua de Hogares. Seleccionada la muestra de segmentos o UPM para estas encuestas, se procedió a levantar el Marco de Listas o Registro de Edificios y Viviendas (REV).

El REV consiste de un listado de todas las estructura comprendidas dentro de los segmentos seleccionados acompañado de su respectivo croquis o mapa que muestra la ubicación de esas estructuras. Este marco se

levantó en fecha lo más próxima a la encuesta con el fin de que estuviera la más actualizado posible. Elaborado el REV se procedió a la selección de la muestra de viviendas según el diseño muestral propuesto.

Este Marco Muestral de Viviendas, que comprende el marco de áreas y el marco de listas, se utilizó hasta 1997 fecha en que se consideró necesario elaborar un nuevo marco dado los problemas de desactualización que presentaba.

II. Marco Muestral de Viviendas 1998

Los Marco Muestrales normalmente se construyen después de finalizar la enumeración censal. En nuestro país los Censos Nacionales programados para 1995 fueron suspendidos debido a la falta de recursos y como consecuencia de los diversos fenómenos naturales que afectaron el país en esa época. A 1998 no se tenía a corto plazo una fecha definida para la realización de censos, y el MMV con 12 años de haberse construido presentaba serios problemas de desactualización que se reflejaban en los resultados de la EHPM. En los últimos años la EHPM venía mostrando un alto crecimiento de la población en la zona rural cuando es de esperar lo contrario. La actualización cartográfica llevada a cabo en el período 1993-1997, en la que se calificó los segmentos por grado de urbanización utilizando los mismos criterios con los que se construyó el marco anterior, mostró fuertes cambios específicamente en la periferia urbana y el rural concentrado.

De esta manera en 1998 con una cartografía actualizada para los censos nacionales y un conteo de viviendas por segmento se iniciaron las actividades para la construcción de un nuevo MMV y del rediseño de la EHPM. El MMV se construyó con base en información proveniente de la actualización cartográfica realizada durante el período 1993-1997, como parte de las actividades preparatorias de los Censos Nacionales de Población y Vivienda. Se utilizó una nueva delimitación de segmentos censales de zonas urbanas y rurales, e incluye el conteo del número de viviendas que se hizo en cada uno de estos segmentos.

Para la estratificación se utilizaron variables geográficas: región, provincia, cantón, distrito y grado de urbanización. Este último se asignó siguiendo los mismos criterios del marco anterior. Dado que no se disponía de información censal reciente, la estratificación por nivel socioeconómico de los segmentos se realizó con base a una calificación visual de las viviendas y su entorno. En este sentido se definieron los criterios, se confeccionaron los instrumentos necesarios y se capacitó a un grupo de funcionarios con la idea de efectuar una calificación visual lo más objetiva posible. De esta forma se clasificó cada segmento censal en uno de cinco grupos: bajo, medio-bajo, medio, medio-alto, alto y rural. Posterior a la construcción del marco se seleccionó la muestra de segmentos o UPM de acuerdo al rediseño propuesto y para cada segmento seleccionado se levantó el REV. Finalmente el MMV quedó constituido por 16 107 segmentos a nivel nacional (cuadro N° 1).

CUADRO N° 1
DISTRIBUCIÓN DE SEGMENTOS Y VIVIENDAS DEL NUEVO MARCO MUESTRAL DE VIVIENDAS POR ZONA URBANA Y RURAL SEGÚN REGIÓN DE PLANIFICACIÓN

REGIÓN	TOTAL		ZONA			
			URBANA		RURAL	
	Segmentos	Viviendas	Segmentos	Viviendas	Segmentos	Viviendas
COSTA RICA	**16.107**	**826.596**	**7.781**	**463.594**	**8.326**	**363.002**
Central	9.669	525.076	6.105	363.854	3.564	161.222
Chorotega	1.419	67.722	387	23.593	1.032	44.129
Pac. Central	992	50.135	461	26.359	531	23.776
Brunca	1.453	64.716	273	16.653	1.180	48.063
Huetar Atlántica	1.669	78.349	439	25.936	1.230	52.413
Huetar Norte	905	40.598	116	7.199	789	33.399

Este marco se ha utilizado para las encuestas de hogares de 1999 al 2002. Con la realización de los Censos de Población y Vivienda del 2000 se está en las actividades para la construcción de un nuevo MMV.

III. Diseño de la Muestra

En la definición del Diseño Muestral de la Encuesta de Hogares de Propósitos Múltiples (EHPM) se contó con la asesoría del consultor Dr. Anis Kumar Maitra (OIT) y del Dr. Fernando Medina (CEPAL) para evaluar el anterior diseño muestral y proponer ajustes de acuerdo a los requerimientos actuales de la encuesta. Uno de estos requerimientos es la inclusión de módulos sobre temas relacionados con el trabajo infanto-juvenil.

3.1 Población y cobertura

La población que investiga la EHPM está definida como el conjunto de todas las viviendas individuales y sus ocupantes, **residentes permanentes** en esas viviendas, tanto del área urbana como del área rural de Costa Rica. Se excluye del estudio a la población residente en las viviendas colectivas (hoteles, hospitales, asilos, cárceles, conventos, etc.).

El ámbito geográfico del MMV en principio se extiende a todo el territorio nacional, sin embargo, por razones de costo se excluyó para la encuesta un grupo de segmentos censales por considerarse de baja densidad poblacional: segmentos de menos de 15 viviendas. Esto significa que la población muestreada es un subconjunto de la población objetivo lo que podría tener como resultado un sesgo en las estimaciones. Como los segmentos excluidos representan un 0.28% del total de segmentos del marco se consideró que el sesgo es poco significativo y por otro lado representa un ahorro importante de recursos para la encuesta.

3.2 Dominios de estudio

Para el rediseño de la muestra de la encuesta, se definieron los mismos dominios de estudio del programa de encuestas de 1987-1998, esto con el fin de mantener la comparabilidad de la serie de datos. Los dominios de estudio geográficos propuestos para la encuesta de 1999 son:

1. Nacional
2. Nacional Urbano y Nacional Rural
3. Región Central: Urbano y Rural
4. Resto de Regiones: Urbano y Rural

3.3 Unidades de muestreo

La EHPM el diseño muestral incorpora dos etapas de selección y por tanto dos diferentes unidades de muestreo.

- **Unidad Primaria de Muestreo (UPM):** en la primera etapa de selección la unidad de muestreo corresponde al segmento censal del MMV.
- **Unidad Secundaria de Muestreo (USM):** en la segunda etapa de selección la unidad de muestreo corresponde a la vivienda y esta es además la unidad última de selección.

3.4 Definición de la muestra

La muestra de la EHPM corresponde a una muestra probabilística de áreas, estratificada y bietápica. Es probabilística porque en cada una de las etapas se utilizan métodos aleatorios de selección que permiten hacer inferencias a la población con base en la teoría de probabilidades. Es de áreas porque la probabilidad de selección está asociada a áreas geográficas dentro de los dominios de estudio. Es estratificada porque previamente a la selección la población se dividió en doce estratos -cada región de planificación dividida por

zona urbana y zona rural- con la finalidad de mejorar su representatividad y aumentar así la eficiencia relativa del diseño muestral. Es bietápica porque en una primera etapa se selecciona segmentos censales (UPM) y en una segunda etapa viviendas (USM) dentro de las UPM seleccionadas en primera etapa.

3.4.1 Determinación del tamaño de muestra

Antes de determinar el tamaño de muestra del diseño revisado, el asesor tomó en consideración las recomendaciones que sobre el tópico de tamaños de muestras se dan en el manual sobre Encuestas de Población Económicamente Activa de la OIT [1].

Para determinar el tamaño de muestra se utiliza la fórmula del muestreo aleatorio simple ajustada por el efecto del diseño. , a saber:

$$n = \frac{Z^2_{(\alpha)}\ \sigma^2}{E^2} * \text{Deff}$$

donde:

$Z_{(\alpha)}$: es el valor crítico de la curva normal estándar para un intervalo de confianza del $(1-\alpha)$ por ciento.

σ^2 : es la variancia de la variable de interés. En el caso de una proporción esta variancia es igual a $P(1-P)$.

E : es el margen de error propuesto como aceptable para estimar la variable de interés.

Deff: es el valor del efecto del diseño obtenido de encuestas anteriores y representa el aumento en la variancia al utilizar diseños muestrales complejos en lugar de una muestreo aleatorio simple.

La encuesta más reciente (1997), da una estimación de la tasa de desempleo de alrededor del 5% para la mayoría de las regiones: $P=0.05$ y $\sigma^2 = P(1-P) = 0.05 * 0.95 = 0.0475$.

El valor de α es 0.05 y el valor para este nivel de significancia en la tabla de la distribución normal estándar es de $Z_{(\alpha)} = 1.96$. El margen de error propuesto es $E=0.01$, esto significa que si la tasa observada de desempleo es $p=0.05$, el verdadero valor de la tasa de desempleo será estimado con un 95% de confianza en el intervalo (0.04, 0.06).

La región Central tiene un mayor peso sobre las estimaciones totales y es para esta región que se dan estimaciones por zona urbana y rural, en el resto de las regiones se agregan los resultados de zona urbana y rural. Siguiendo el criterio del asesor de utilizar el efecto del diseño más alto para la tasa de desempleo se tiene que el mayor valor del efecto del diseño se da en la zona urbana de la región Central cercano a 2.45.

De los datos de la reciente encuesta se ha encontrado que hay un promedio de 1.5 personas económicamente activas por vivienda y cerca de 1.1 hogares por vivienda en el país.

Con la información anterior se tiene:

$$n = \frac{(1.96)^2\ (0.0475)}{(0.01\ )^2} * 2.45$$

[1] *OIT. Manual sobre Encuestas de la Población Económicamente Activa. Página 221.*

=	1,824.75 * 2.45	
=	4,470.662	
=	4,471	personas económicamente activas (y dividido por 1.5)
=	2,980.667	hogares (y dividido por 1.1).
=	2,709.697 ≅ 2710	viviendas.

El diseño muestral considera cuatro estratos importantes de estimación: Central Urbano, Central Rural, Resto del País Urbano y Resto del País Rural. El primer supuesto es que se requiere aproximadamente el mismo grado de precisión para las estimaciones de cada uno de estos cuatro subniveles nacionales. Sobre la base de este supuesto, el tamaño de muestra total requerido para el total país es de:

$$n = 4 * 2{,}710 = 10{,}840 \text{ viviendas.}$$

La tasa de respuesta de la EHPM se ha mantenido alrededor del 90%, así el tamaño de muestra efectivo que se espera obtener es de 9,756 viviendas, o 10,732 hogares. El asesor recomendó que el tamaño de muestra no debería ser incrementado en esta etapa para compensar la no respuesta y hacer los esfuerzos necesarios para que se mantenga a estos niveles.

3.4.2 Distribución de la muestra

Costa Rica y otros países de la región se han unido al programa de erradicación del trabajo infantil, pasando a ser esta otra área de gran interés en la encuesta. Por este motivo en el diseño muestral se ha tomado en consideración esta variable para la afijación o distribución de la muestra de modo que se garantice para esta variable una buena estimación a nivel nacional y por zona urbana y rural.

Se ha observado que existen algunas diferencias apreciables en el error estándar y en las tasas de participación económica del menor trabajador entre la zona urbana y rural, por ello, la distribución de la muestra toma en consideración esta variación. En el cuadro N° 2 se presenta la afijación que se realizó de la muestra a partir de la distribución del número de viviendas en la zona urbana y rural obtenida de la actualización cartográfica y de los datos sobre trabajo infantil de la encuesta de julio de 1995.

CUADRO N° 2
AFIJACIÓN DE LA MUESTRA POR ZONA URBANA Y RURAL A PARTIR DEL ERROR ESTÁNDAR DE LA PARTICIPACIÓN ECONÓMICA DEL MENOR TRABAJADOR

Zona	Viviendas N_i	Error Estándar (σ)	$N_i * \sigma_i$	$\frac{N_i * \sigma_i}{\Sigma N_i * \sigma_i}$	n_i (5) * n
(1)	(2)	(3)	(4)	(5)	(6)
Total	826.596	-	5.221,986	1,000	10.840
Urbana	463.594	0,005	2.317,970	0,444	4.812
Rural	363.002	0,008	2.904,016	0,556	6.028

donde:

N_i : número de viviendas en el MMV en la zona i, para i=1, 2

σ : error estándar de la participación económica del menor trabajador.

n : muestra total de viviendas a seleccionar.

n_i : muestra de viviendas a seleccionar por zona urbana y rural.

Lo anterior indica que se necesitan 4,812 viviendas en la zona urbana y 6,028 viviendas en la zona rural para dar estimaciones bastantes confiables del trabajo infantil a nivel urbano y rural.

En la EHPM se ha observado para la tasa de desempleo coeficientes de homogeneidad bastante bajos. Además, debido a los costos que implica un mayor número de UPM en la muestra (actualización, trabajo de campo, proceso, etc.), se recomendó, para determinar la muestra de UPM, mantener el promedio de 15 entrevistas por UPM que actualmente se realizan en la encuesta (cuadro N° 3).

CUADRO N° 3
NÚMERO DE UPM A SELECCIONAR POR ZONA URBANA Y RURAL

Zona	n_i	USM por UPM	Número de UPM
(1)	(6)	(7)	(6) / (7)
Total	10.840		724
Urbana	4.812	15	322
Rural	6.028	15	402

Como las seis regiones de planificación son subpoblaciones importantes de estudio y se requieren estimaciones confiables a este nivel, se consideró importante distribuir la muestra de segmentos por zona urbana y rural entre las seis regiones de planificación. En este sentido, la distribución de la muestra de segmentos se realizó tomando en cuenta la variación relativa, coeficiente de variación, en la estimación del número de desempleados. Con el fin de suavizar las variaciones de un año a otro se tomó la variación relativa promedio de esta variable entre los años 1995, 1996 y 1997. La distribución final de la muestra de segmentos en cada una de las regiones por zona urbana y rural se presenta en el cuadro N° 4

CUADRO N° 4
DISTRIBUCIÓN DE LA MUESTRA DE UPM POR ZONA SEGÚN REGIÓN

REGIÓN	UPM	Variación Relativa	N_i*V_i	N_i*V_i	n_i
Total	**7.786**		**1.116,30**		**322**
Central	6.110	0,1149	702,04	0,6289	202
Chorotega	387	0,2478	95,90	0,0859	28
Pac. Central	461	0,2324	107,14	0,0960	31
Brunca	273	0,3039	82,96	0,0743	24
H. Atlántica	439	0,2024	88,85	0,0796	26
H. Norte	116	0,3397	39,41	0,0353	11

3.5 Selección de la muestra

Para la selección de la muestra, se dividió el MMV en doce submarcos, correspondientes a las seis regiones de planificación divididas por zona urbana y rural. El procedimiento de selección dentro de cada submarco se realizó dando un ordenamiento de las UPM por tamaño (número de viviendas) de menor a mayor.

3.5.1 Selección de Unidades Primarias de Muestreo (UPM)

La selección de la muestra de UPM sobre cada submarco se efectuó en forma sistemática después de un ordenamiento según el número de viviendas por UPM. El criterio de hacer una selección sistemática de una lista ordenada por tamaño, se debe a que el ordenamiento por tamaño de las unidades produce una

estratificación natural, lo cual garantiza una mejor representatividad de la muestra bajo el supuesto de que las UPM con tamaño similar presentan características bastante homogéneas.

La selección sistemática da igual probabilidad de selección para cada Unidad del Marco, por tanto la probabilidad de selección de cada UPM está dada por $\mathbf{P_{hi} = 1 / N_{hi}}$, donde P_{hi} es la probabilidad de selección y N_{hi} es el número de UPM en la i-ésima zona de la h-ésima región (para i=1, 2 y h=1, 2, 3, 4, 5, 6).

3.5.2 Selección de Unidades Secundarias de Muestreo (USM)

Después de la selección de la muestra de UPM en la primera etapa, se procedió al levantamiento del marco de segunda etapa. Este marco corresponde a un Marco de Lista llamado Registro de Edificios y Viviendas (REV) que tiene como objetivo identificar, registrar y describir el uso o destino de los diferentes tipos de edificios y viviendas, así como levantar el listado de jefes de hogar para aquellas estructuras dedicadas a viviendas dentro de cada una de las UPM seleccionadas en la primera etapa. Una vez digitada y procesada esta información se pasó a la etapa de selección de las viviendas o USM.

Antes de seleccionar las USM se tomó la decisión de que la muestra fuera autoponderada por región y zona, esto significa que todas las viviendas dentro de una misma región de planificación y zona tienen la misma fracción de muestreo. En este sentido, para mantener la autoponderación, se fijó en un cuarto la fracción de muestreo en segunda etapa para los segmento urbanos y una fracción de muestreo de un tercio en la segunda etapa para los segmentos rurales.

La selección de USM se efectúa de forma sistemática en conglomerados de tres viviendas con probabilidades iguales. Aunque la selección de viviendas en conglomerados compactos tiende a aumentar el grado de homogeneidad dentro de las UPM, se espera que este aumento no sea muy importante y se prefiere a una selección aleatoria o sistemática de viviendas individuales por las ventajas que implica en cuanto al recorrido dentro de las UPM y a una mejor labor de supervisión en el trabajo de campo. Debido a que se selecciona una fracción fija en segunda etapa, el número final de USM dependerá del tamaño de las UPM.

3.6 Rotación de la muestra

No se ha planteado dentro de los objetivos de la encuesta hacer estimaciones de cambio entre dos períodos o estimaciones de promedios para dos períodos, para las cuales se requiere un traslape de la muestra o muestras independientes respectivamente. En este sentido el esquema de rotación que se efectúa es de la muestra de viviendas en un 25% de las UPM, se hace únicamente con el fin de evitar el cansancio de los entrevistados y por consiguiente disminuir los posibles rechazos que se puedan presentar por este motivo.

3.7 Factor de expansión

El factor de expansión básico para cada vivienda seleccionada es determinado por el diseño de la muestra y en un diseño bietápico, como es el caso de la EHPM, este factor es calculado al multiplicar el inverso de las probabilidades de selección de cada etapa:

$$F_{hij} = 1 / f_{hi\,1} * 1 / f_{hij\,2}$$

donde

F_{hij} : es el factor de expansión básico para la h-ésima región, la i-ésima zona en la j-ésima UPM.

$f_{hi\,1}$: fracción de muestreo en la primera etapa de selección para la h-ésima región y la i-ésima zona.

$f_{hij\,2}$: fracción de muestreo en la segunda etapa de selección para la h-ésima región, la i-ésima zona en la j-ésima UPM.

3.8 Ajustes al factor de expansión

Es común que se presenten diferencias entre lo planificado en oficina y lo observado en el campo, debido a esto el factor de expansión básico debe ser ajustado para mantener las probabilidades iniciales de selección. Estas diferencias que afectan las probabilidades de selección se deben principalmente a dos aspectos: a diferencias en el tamaño de las UPM en cuanto al número de viviendas (actualización), y a diferencias en cuanto al número de viviendas seleccionadas y viviendas realmente encuestadas (tasa de respuesta).

Ajuste por actualización.

La población del país no permanece estática, experimenta cambios tanto en su estructura demográfica como en sus características socioeconómicas, pero la población a la que se refiere el MMV se refiere a la población en un momento específico, no refleja esos cambios que se dan año a año. Esto implica, otro que las probabilidades iniciales de selección de la muestra se van desactualizando y se va perdiendo confiabilidad en las estimaciones que brinda la encuesta. Para solventar este problema se hace necesario actualizar el dato poblacional del marco de forma tal que se vaya incorporando ese dinamismo poblacional.

La actualización del dato poblacional del MMV incluye, por una parte, una continua revisión de los tamaños de las UPM con la consiguiente alteración de los factores de expansión, por otra parte, una nueva selección de las USM incorporando así los cambios sociodemográficos que se dan en la UPM. En este sentido el asesor de la OIT planteó dos formas de actualización de las UPM para incorporar estos cambios:

- **Total de la UPM:** el supervisor debe registrar el número de viviendas no listadas, ya sea por resultado de una mala actualización cartográfica o por haber sido construidas después de la actualización.
- **Dentro de las USM:** el supervisor debe registrar el número de viviendas no listadas en los conglomerados seleccionados dentro de la UPM. En este caso, una estimación del número de nuevas viviendas en la totalidad de la UPM se obtiene multiplicando el número de nuevas viviendas por el inverso de la fracción de muestreo dentro de las UPM.

Ambas formas de actualización pueden utilizarse en diferentes UPM, pero si es importante indicar cuál procedimiento se utilizó para el ajuste final del número de nuevas viviendas. Por ejemplo, si para un estrato se tienen N_{hi} UPM en el marco y n_{hi} UPM en la muestra, y de las n_{hi} UPM se actualizan $n_{hi\,1}$ completamente y las restantes $n_{hi\,2}$ se actualizan sólo en los conglomerados seleccionados, el número estimado de nuevas viviendas (m'_{hi}) en las n_{hi} UPM de la muestra es:

$$\mathbf{m'_{hi} = m'_{hi1} + (m'_{hi2} * F_{hi2})}$$

donde:

m'_{hi} : número de nuevas viviendas en la muestra de la h-ésima región y de la i-ésima zona.

m'_{hi1} : número de nuevas viviendas encontradas en las n_1 UPM actualizadas completamente en la h-ésima región y en la i-ésima zona. .

m'_{hi2} : número de nuevas viviendas encontradas en las n_2 UPM actualizadas solo en los conglomerados seleccionados de la h-ésima región y de la i-ésima zona.

$F_{hi\,2}$: inverso de la fracción de muestreo de segunda etapa para la h-ésima región y la i-ésima zona.

El número de nuevas viviendas estimado en el marco está dado por:

$$M'_{hi} = \frac{N_{hi}}{n_{hi}} * m'_{hi}$$

con lo que el total actualizado de viviendas en el marco es:

$$M_{hi}^{*} = M_{hi} + M'_{hi}$$

donde

M'_{hi} : es el número estimado de nuevas viviendas en el marco para la h-ésima región y la i-ésima zona.

N_{hi} : es el número de segmentos en el marco en la h-ésima región y la i-ésima zona.

n_{hi} : es el número de segmentos en la muestra en la h-ésima región y la i-ésima zona.

m'_{hi} : es el número estimado de nuevas viviendas en la muestra en la h-ésima región y la i-ésima zona.

M_{hi} : es el número actual de viviendas registradas en el marco para la h-ésima región y la i-ésima zona.

M_{hi}^* : es el número actualizado de viviendas en el marco para la h-ésima región y la i-ésima zona.

M_{hi}^*/ M_{hi} : es el factor de ajuste por el cual es multiplicado el factor de expansión inicial.

El factor básico de expansión es ajustado según los resultados anteriores. Es importante recordar que el factor de expansión puede ser determinado solo después del trabajo de campo, porque tanto M^*_{hi}, M'_{hij} y m'_{hi} dependen del número de viviendas no listadas estimadas en la muestra.

Ajuste por tasa de respuesta

Es necesario ajustar los factores básicos de expansión para tomar en consideración las viviendas seleccionadas pero no encuestadas con la finalidad de eliminar el sesgo debido a la no entrevista. Este ajuste a los factores básicos de expansión se efectúa a nivel de zona dentro de cada región de planificación:

$$TRV_{hi} = \frac{v_{hi(e)}}{v_{hi(e)} + v_{hi(ne)}} = \frac{v_{hi(e)}}{v_{hi}}$$

donde

$v_{hi\,(e)}$: total de viviendas seleccionadas y entrevistadas en la h-ésima región y la i-ésima zona.

$v_{hi\,(ne)}$: total de viviendas seleccionadas y no encuestadas en la h-ésima región y la i-ésima zona.

v_{hi} : total de viviendas seleccionadas en la h-ésima región y la i-ésima zona.

Los factores que influyen en la no-respuesta y que impiden la realización de las entrevistas merecen un análisis adicional buscando alternativas para su tratamiento y posibles ajustes.

Ajuste por proyecciones de población

Los estimadores naturales incorporan en el factor de expansión las características propias del diseño muestral, pero se ven afectados por los problemas identificados en el marco. Una forma de solucionar estos problemas es el uso de estimadores que incorporan fuentes externas a la muestra como una compensación a los problemas detectados: envejecimiento del marco muestral, desconocimiento del total de viviendas en la población, falta de capacidad de la muestra para captar una variación en el tamaño de las UPM o desconocimiento de la precisión de los listados de viviendas.

Dado que el programa de EHPM no dispone de un marco de nuevas construcciones que compensa la desactualización del MMV, el diseño muestral de la encuesta contempla un ajuste según las proyecciones oficiales de población2, lo cual permite compensar las variaciones del marco muestral producto del crecimiento poblacional.

$$T_h = \frac{y'_{PHh}(t)}{y'_{Hh}(t)}$$

donde:

$y'_{PHh}(t)$: población estimada que se espera obtener en el año t de la encuesta según tasa de crecimiento de proyecciones de población para la h-ésima región.

$y'_{Hh}(t)$: población estimada de la encuesta sin ajuste de proyección para la h-ésima región.

El factor de expansión final resulta de multiplicar el factor de expansión inicial por los diferentes ajustes:

$$F_{hi}^{*} = = F_{hi} * \frac{M_{hi}^{*}}{M_{hi}} * \frac{1}{TRV_{hi}} * T_h =$$

El factor de expansión final ajustado por proyecciones se aplica a todas las unidades (viviendas o personas) dentro de cada zona y región de planificación, luego se van agregando los datos para obtener las estimaciones para cada subpoblación de interés.

3. 9 Cálculo de errores muestrales

Para el cálculo de errores muestrales de la EHPM se utilizará el módulo CENVAR [3] del programa estadístico IMPS (Integrated Microcomputer Processing System). El CENVAR produce una tabla de salida o reporte con los siguientes indicadores:

1. El valor estimado del parámetro.
2. El error estándar.
3. El coeficiente de variación (C.V.).
4. El intervalo de confianza del 95 por ciento.
5. El efecto del diseño (Deff).
6. El número de observaciones (n).

Finalmente, los principales resultados de la EHPM se publican junto con algunos errores muestrales para las variables más importantes de la encuesta como son las tasas de ocupación, desocupación y pobreza.

2/ Centro Latinoamericano de Demografía, CELADE. Estimaciones y Proyecciones Regionales de Población. San José.1988.
3/CENVAR: Variance Calculation System.1995.

MARCOS DE ÁREAS

OFICINA NACIONAL DE ESTADÍSTICA
CUBA

ÍNDICE

Página

Tema 1 Los Censos de Población y Vivienda y los Marcos de Muestreo Utilizados en las Encuestas de Hogares 169
1.1 Relación entre la categoría censal y el marco de muestreo utilizado para la selección de las viviendas 169
1.2 Utilización de material censal para la construcción de los marcos de muestreo 170
1.3 Preparación y adaptación de los mapas censales para la elaboración del marco de muestreo 171
Tema 2 Diseño y Construcción del Marco de Muestreo 171
2.1 Elementos generales sobre un marco muestral 171
2.2 Importancia de los marcos de áreas para la selección de viviendas 172
2.3 Requerimientos de información, dominios de estudio e inferencia estadística 173
2.4 Los marcos maestros de muestreo y los diseños de muestras maestras 173
2.5 Procedimientos para la formación de las unidades del marco de muestreo 174
Tema 3 Aspectos Estadísticos Asociados a la Construcción del Marco de Muestreo 175
3.1 Procedimientos para la estratificación del marco de muestreo 175
3.2 Selección de l as unidades de muestreo y requerimientos de información 175
3.3 Métodos para la asignación de la muestra 175
3.4 Procedimientos para la selección de la muestra 176
3.5 Diseños en tres etapas: ventajas y desventajas 176

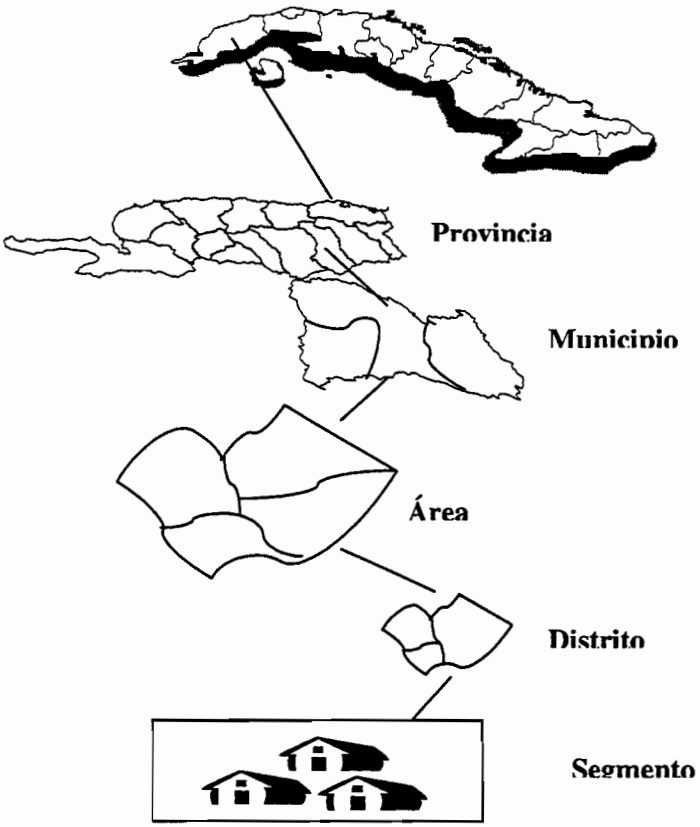
TEMA 1. Los Censos de Población y Vivienda y los Marcos de Muestreo Utilizados en las Encuestas de Hogares

1.1 Relación entre la cartografía censal y el marco de muestreo utilizado para la selección de las viviendas

Entre las principales tareas que desarrollan los países que periódicamente acostumbran a hacer censos, se encuentra la organización de la cartografía censal. La misma permite disponer de un completo y pormenorizado archivo de mapas y planos del país. Dentro de este conjunto de planos y mapas, revisten principal importancia aquellos pertenecientes a los lugares habitados tanto urbanos como rurales.

El objetivo principal de la cartografía censal, es poder disponer de mapas de todas las unidades políticos-administrativas y censales del país. Este material debe, en primer término, servir de base para la organización del dispositivo censal y una vez dentro de la etapa de toma de información, servirá para controlar y supervisar las labores de levantamiento. Por otro lado constituye un valioso instrumento para evitar que se produzcan duplicidades u omisiones en el terreno y para realizar comparaciones con censos anteriores.

Los mapas deben contener de manera clara los límites provinciales, municipales y de manera general, los límites administrativos censales (áreas, distritos, segmentos). El gráfico siguiente, presenta la subdivisión en unidades territoriales, adoptada por la cartografía censal para el Censo de Población y Viviendas de Cuba del año 2002.



Los mapas y planos deben permitir identificar claramente los límites territoriales de cada unidad en cuestión. En la parte urbana deben estar indicados con todo detalle las manzanas, calles y otras referencias

locales, mientras que en la parte rural se deben reflejar claramente los accidentes geográficos, carreteras, caminos, ferrocarriles y otras referencias, que le permitan al personal de campo, organizar su recorrido, definir exactamente las áreas que debe enumerar, organizar las labores de supervisión, etc.

Para confeccionar estos mapas, resulta imprescindible realizar un peine preliminar del terreno y además, lo que se ha dado llamar el Registro Previo que no es más que un listado de las viviendas que existen en el terreno y los ocupantes de las mismas.

En el caso de las encuestas por muestreo, resulta especialmente importante la cartografía, ya que el mapa sirve para señalar la ubicación de las unidades a investigar, su distribución en el terreno y otros detalles útiles para organizar el levantamiento de la encuesta en cuestión.

1.2 Utilización del material censal para la construcción de los marcos de muestreo

El último Censo de Población y Viviendas se realizó en Cuba en el año 1981. Por diferentes razones de índole fundamentalmente económica, no fue posible ejecutar esta importante investigación estadística en la década de los noventa del siglo XX.

Actualmente está en curso la etapa preparatoria del Censo de Población y Viviendas a levantarse en Septiembre del 2002. Para ello, se han aprovechado las experiencias de los Censos de 1970 y 1981, y los preparativos del Censo del año 1990 (que no se llevó a efecto), con vistas a lograr, desde el punto de vista metodológico y organizativo, la mayor eficiencia en los resultados y en la economía de la investigación.

Dentro de estas actividades se encuentra la organización y perfeccionamiento de la cartografía censal, de vital importancia para garantizar la orientación en el terreno del personal de campo.

A pesar de que el país no ha contado con un censo actualizado en los últimos 20 años, esto no ha sido obstáculo para que desde el año1982, se haya iniciado la organización de lo que en su momento se llamó Marco Muestral Maestro, estructurado según la División Político Administrativa (DPA) del país cuyas Unidades de Primera Etapa de selección estuvieron constituidas por los Distritos Censales.

En ese año se elaboró el Diseño Muestral General (GMG) también llamado Muestra Maestra (MM) que serviría de base al Sistema de Encuestas de Hogares y quedaría organizado en tres bloques, los dos primeros para el desarrollo de las encuestas de carácter continuo y el tercero para el desarrollo de encuestas eventuales.

Los trabajos de conformación del marco y selección de las unidades muestrales se realizaron en el año 1983, instrumentándose la selección de forma automatizada mediante máquinas computadoras. En diciembre de 1983 se concluyó desde el punto de vista técnico el diseño del Bloque II referido a la Encuesta Nacional de Ocupación (ENO) implementándose además en computadora las diferentes etapas de muestreo y la confección de las guías de visita.

A fines del I Semestre de 1984 se obtuvo por computadora el listado de viviendas seleccionadas que intervendría durante todo el primer año de ejecución de la ENO y se trabajó en la elaboración de los bloques correspondientes a la Encuesta de Presupuesto Familiar y el otro destinado a Encuestas Eventuales.

En el año 1990, ante las dificultades económicas, materiales y energéticas del país en los comienzos del período especial, se decidió transformar el Muestreo por Listas que se venía aplicando, por un Muestreo por Áreas para garantizar la continuidad del DMG y con ello la selección de las muestras de las encuestas a la población.

En el año 1995 se definió una nueva división territorial por debajo del municipio denominada Área Geográfica Muestral (AGEM), la que vino a sustituir los Distritos censales que se venían utilizando hasta ese momento. Los principales elementos en que se basó la definición de esta nueva división territorial fueron, en la parte Rural, las llamadas Unidades Básicas de Información Territorial (UBIT) definidas por el Instituto de Planificación Física y en la parte Urbana, los Distritos Censales.

El DMG conformado por las AGEM, comenzó por un Muestreo de Listas y luego en el año 1998 se inició el proceso de transición del muestreo por lista al muestreo por áreas.

En 1995, sólo se logró la división de planos municipales en AGEM y se confeccionaron los expedientes cartográficos de las AGEM seleccionadas para participar en la Encuesta Nacional de Migraciones Internas de Cuba.

No es hasta el período Septiembre 1998 - Junio 1999, que se dispone de los recursos necesarios para confeccionar los expedientes cartográficos de todas las AGEM del país. Los mismos contienen los mapas o planos, derroteros y los listados de las viviendas contenidas en las AGEM, confeccionados a partir de la actualización en el terreno del Marco Muestral Maestro.

Actualmente, a partir de las labores preparatorias que se desarrollan para el próximo levantamiento del Censo de Población y Viviendas en Septiembre del 2002, se ha realizado una revisión exhaustiva de todos los elementos que conforman el Marco Muestral Maestro y a partir de la revisión y ajuste de las AGEM se han construido los actuales Distritos Censales.

Las cantidades promedio de viviendas que contienen las diferentes unidades territoriales del Censo aparecen en el siguiente cuadro:

	URBANO	RURAL
Segmento	85	65
Distrito	255	195

La estructura censal del país después de la última actualización de mayo del 2002 es la siguiente:

	URBANO	RURAL	TOTAL
Asentamientos	657	6325	6982
Distritos	10765	5094	15859
Segmentos	32935	14226	47161

1.3 Preparación y adaptación de los mapas censales para la elaboración del marco de muestreo

Para la preparación de los mapas censales para la elaboración de Marco de requerirán, en la parte rural, de hojas cartográficas de 1 en 10 000 de las unidades muestrales de primera etapa seleccionadas y planos de 1 en 2 500, 1 en 3 000 o 1 en 5 000 de los poblados o asentamientos con características urbanas. Con esos expedientes y los planos o mapas, se requiere recorrer el área comprendida por las Unidades de Muestreo fijando los límites para su división.

En la parte urbana la manzana se divide en grupos de cinco viviendas en promedio para conformar las secciones, de tal manera que las secciones cubran todo el terreno, incluidos los espacios no utilizados, dentro de alguna sección colindante.

En la parte rural se requiere ir agrupando las viviendas más próximas en grupos de cinco viviendas en promedio y que queden dentro de límites que se puedan describir para definir el área.

TEMA 2. Diseño y Construcción del Marco de Muestreo

2.1 Elementos generales sobre un marco muestral

Antes de iniciar un acercamiento al tema de Marco Muestral Maestro en Cuba, resulta conveniente definir en primer lugar qué es el Marco Muestral Maestro (MMM).

Se llama **Marco Muestral Maestro** (MMM) al conjunto de unidades a partir del cual puede ser seleccionada una muestra. A partir de aquí se le denominará simplemente **Marco**.

Desde el punto de vista de su contenido, la estructuración del Marco, contempla desde la formación del listado de unidades a muestrear, es decir, de los elementos integrantes de una población dada, hasta la información necesaria que permita su identificación y clasificación para mejorar los procesos de selección, estratificación o estimación a partir de una muestra. Otra información valiosa que deben contener los marcos, la constituye las medidas de tamaño de las diferentes subdivisiones, también conocidas como unidades de muestreo, que los conforman.

Los marcos se utilizan en los períodos intercensales para, a partir de una muestra de elementos de una población dada, realizar inferencias de la misma.

Los **elementos de una población** son las unidades elementales sobre las que se busca información y de las cuales se hacen inferencias.

Se denomina **Población** al conjunto de unidades del que se requiere información, para obtenerla es necesario medir o contar en cada unidad determinadas características o clasificar las unidades integrantes de acuerdo a ciertos parámetros.

Dentro de Sistema de Encuestas de Hogares de Cuba, la población está conformada en unos casos por las viviendas particulares y los hogares que residen en ellas como es el caso de la Encuesta sobre la Situación Económica de los Hogares y en otros, por las personas que son residentes permanentes de las viviendas particulares, como es el caso de la Encuesta Nacional de Ocupación, Encuesta Nacional de Factores de Riesgo para la Salud, Encuesta sobre el VIH/SIDA, Encuesta sobre la Salud, Bienestar y Envejecimiento del Adulto Mayor, etc.

Anteriormente se hizo referencia al término **Muestra**. Casi siempre, en la mayoría de las investigaciones No es factible obtener información de todas las unidades de la población, o lo que es lo mismos, realizar un censo, por la gran cantidad de recursos materiales y humanos que se requiere o por lo complejo de los estudios. En estos casos, las observaciones se limitan a un subconjunto de las unidades de la población que es lo que se denomina **Muestra**.

Para seleccionar las unidades que conforman la Muestra es necesario disponer de información con relación al conjunto de unidades que conforman la población en estudio y de algunos parámetros de clasificación de cada una de esas unidades. Es precisamente esa información la que debe estar contenida en el Marco Muestral.

Las unidades de muestreo pueden contener un solo elemento, varios o ninguno. Cuando las unidades de muestreo contienen varios elementos, se les denomina **conglomerados**, como puede ser una manzana en una ciudad. En la práctica de encuestas es muy frecuente dentro de los conglomerados realizar lo que llamamos submuestreo de unidades o de conglomerados en etapas sucesivas, denominadas Unidad de Primera Etapa de Muestreo (UPM), Unidad de Segunda Etapa de Muestreo (USM), Unidad de Tercera Etapa de Muestreo (UTM), etc.

2.2 Importancia de los marcos de áreas para la selección de viviendas

En sentido estricto el Marco no es más que un listado de unidades, en sentido amplio debe incluir toda la información posible referente a esas unidades la cual puede ser utilizada para la estratificación, clasificación y posterior estimación.

Los marcos de áreas son muy convenientes para llevar a cabo la selección de viviendas ya que permiten asociar las viviendas con un área geográfica determinada, ya sea una manzana, segmento o sección, la cual está debidamente definida por límites físicos.

El hecho de utilizar marcos de áreas permite mantener por un tiempo relativamente largo la identificación de las diferentes unidades de muestreo ya que la clasificación por áreas geográficas es relativamente más estable que la clasificación en listas. Por otro lado se refleja de manera automática el crecimiento natural de la población.

Desde el punto de vista del trabajo de campo tiene la ventaja de que es más fácil y rápido identificar las áreas de trabajo.

El proceso de selección se simplifica mucho, en primer término, porque se concentra la muestra en bloques compactos de viviendas próximas denominados secciones. Además, luego de realizar una primera selección se puede establecer un orden de rotación de las secciones que integran el territorio quedando seleccionadas las viviendas de manera automática.

En el caso particular de Cuba el Marco Muestral está conformado como se describe en el siguiente cuadro

Unidades Muestrales	**URBANO**	**RURAL**
Unidad de Primera Etapa	AGEM	AGEM
Unidad de Segunda Etapa	Manzana	Segmento
Unidad de Tercera Etapa	Sección	Sección

2.3 Requerimientos de información, dominios de estudio e inferencia estadística

El primer requerimiento de información es disponer de mapas y planos actualizados de las AGEM que integran el territorio. Se deben definir las características técnicas de los mapas, las simbologías para representar viviendas y otras edificaciones y los diferentes elementos que se encuentran en el terreno, accidentes geográficos, carreteras, caminos, etc. También es necesario contar con los listados de las viviendas ubicadas en las AGEM. Con estos elementos se conforma el expediente cartográfico.

Los dominios de estudio deben ser determinados de acuerdo al tema de la encuesta en cuestión, un punto de partida puede ser la división político administrativa del país o región que se quiera estudiar. Debe tenerse en cuenta si se va a estudiar la zona urbana o rural y el alcance de la investigación a realizar.

Con relación a la inferencia estadística a realizar a partir de una encuesta por muestreo, es obvio que como toda herramienta que se utilice, las muestras también tienen sus limitaciones. Muestras diferentes conducen a resultados diferentes y mientras mayor sea el conocimiento previo de la población objeto de investigación, tanto mejor será la calidad del diseño.

Por ejemplo una muestra diseñada para dar resultados a nivel nacional, no reproduce la situación a nivel provincial o municipal, lo mismo ocurre si se diseña una muestra por regiones. Si se quieren hacer inferencias sobre una población determinada, se debe tener idea de a qué nivel se necesitan producir tabulados.

También se debe contar con un población lo suficientemente grande que haga racional el uso del muestreo. La inferencia estadística sobre divisiones de la población pequeñas, vuelve incosteable una investigación continua.

2.4 Los marcos maestros de muestreo y los diseños de muestras maestras.

La Muestra Maestra (MM) no es mas que el diseño de una muestra compleja de gran alcance en el espacio y en tiempo, que servirá de base para garantizar las encuestas periódicas con el estudio de sus indicadores y de la dinámica de los mismos. Una Muestra Maestra debe mantener la flexibilidad necesaria para la aplicación de encuestas eventuales sobre submuestras seleccionadas con ese fin.

Los Marcos Maestros de Muestreo deben ser capaces de ofrecer toda la información necesaria para cubrir los requerimiento de una MM y por tanto en su organización deben estar previstas las labores de actualización periódica del mismo las que incluyen el pateo de terreno y la actualización de los expedientes cartográficos de las Unidades Muestrales que comprende.

2.5 Procedimientos para la formación de las unidades del marco de muestreo

Para la formación de las Unidades del Marco de Muestreo se deben tener en cuenta los requerimientos de los usuarios y las experiencias que pudieran existir de los diseños anteriores. Las unidades muestrales no son más que las unidades que conforman el marco sobre el que se realiza la selección muestral.

El Marco Muestral que posee la Oficina Nacional de Estadística, está estructurado según la División Política Administrativa. Los municipios fueron divididos en áreas llamadas Áreas Geográficas Muestrales (AGEM), unidades que tienen en promedio 180 viviendas. Las áreas a su vez fueron divididas en Manzanas en la parte urbana y en Segmentos en la parte rural, ambas fueron divididas en secciones.

Como estratos geográficos se considerarán las provincias y los municipios en que está dividido el territorio, según la División Político Administrativa.

Dentro de cada estrato se realizó una estratificación implícita de las AGEM que comprende dos criterios de clasificación, que se aplican al darle un ordenamiento a las AGEM antes del proceso de selección de la muestra.

Primer criterio: Asentamiento Poblacional.

Se ordenaron las AGEM según los códigos de los asentamientos poblacionales a que pertenecen, que comienzan por la ciudad cabecera municipal con el código 001.

Segundo Criterio: se utilizó la doble clasificación por zona urbana y rural y las características de la distribución de las viviendas que radican en cada AGEM, de los cuales resultaron 8 niveles de clasificación:

Urbano	llano montaña
Rural-Llano	concentrado mixto disperso
Rural-Montaña	concentrado mixto disperso

Las medidas de tamaño promedio de las diferentes unidades muestrales son las siguientes:

Unidades Muestrales	**URBANO**	**VIVIENDAS**	**RURAL**	**VIVIENDAS**
Unidad de Primera Etapa	AGEM	180	AGEM	180
Unidad de Segunda Etapa	Manzana	30	Segmento	90
Unidad de Tercera Etapa	Sección	5	Sección	5

Con respecto a la identificación de límites se recomienda en la parte urbana, que las fronteras de las manzanas seleccionadas sean las calles y en la periferia de las ciudades, los ríos, las líneas de ferrocarril, etc. Las señalizaciones, los letreros, los semáforos pueden servir de puntos de referencias para ubicarnos mas fácilmente en el terreno. En la parte rural, los límites, los cuales pueden ser caminos, carreteras, canales de riego o de drenaje, vías de ferrocarril, ríos, lagos, presas, arboledas, sembrados u otros accidentes físicos reconocibles.

La numeración de las AGEM se realizó dentro de los Municipios/Parte. En la parte urbana, las manzanas se numeraron dentro de las AGEM y la numeración de las secciones se realizó dentro de las manzanas

seleccionadas. De forma similar se realizó en la parte rural, numerándose los Segmentos dentro de las AGEM y las secciones dentro de los Segmentos.

TEMA 3. Aspectos Estadísticos Asociados a la Construcción del Marco de Muestreo

3.1 Procedimientos para la estratificación del marco de muestreo

Se denomina estratificación, a la subdivisión de la población o universo de unidades de acuerdo a determinadas condiciones o criterios en subpoblaciones llamadas **estratos** a los que corresponde una selección independiente de la muestra. Es decir, una vez que la población está dividida en grupos tales que los elementos dentro de cada grupo son más parecidos que los elementos de la población vista como un todo, la muestra se extrae separadamente de cada grupo.

El objetivo de la estratificación es, obtener la mayor precisión en los cálculos y garantizar los niveles de salida de la información.

Los estratos a su vez pueden estar divididos en **subestratos** y esta nueva división se hace con el objetivo de que cada uno de ellos tenga garantizada su representación en la muestra.

El Marco Muestral de Cuba tiene con estratos las provincias y municipios en que está dividido el país de acuerdo a la División Político Administrativa vigente.

Ya se hizo referencia anteriormente a que el objetivo de la estratificación es obtener la mayor precisión en lo cálculos, por lo que en al diseño cubano se incluyen dentro de los estratos geográficos la clasificación en parte urbana y rural.

3.2 Selección de las unidades de muestreo y requerimientos de información

En Cuba la Muestra Maestra tiene en total 4000 AGEM, que equivalen a 40 000 viviendas. Las Unidades de Primera Etapa (AGEM) se seleccionaron con Probabilidad Proporcional al Tamaño dado este por el número de viviendas particulares con residentes permanentes.

El incremento de la Muestra se realizó en dos fases, en 1999 se entrevistaron 20 000 viviendas y en al año 2000 se elevó esta cantidad a 40 000 para aumentar la precisión de los resultados. El solapamiento de la muestra es de un 50 %

Las Unidades de Segunda Etapa parte urbana: las Manzanas o Unidades de manzana (de 30 viviendas o más) y en la parte rural los Segmentos (con 90 viviendas o más), se seleccionaron con Probabilidad Proporcional al Tamaño dado este por el número de viviendas particulares con residentes permanentes,

Las Unidades de Tercera Etapa o Secciones se seleccionarán de forma sistemática con Probabilidades Iguales.

Dando lugar este diseño, a una muestra autoponderada a nivel de Municipio/Parte, que es aquella muestra cuyas unidades elementales tienen la misma probabilidad de aparecer en la muestra, dentro del Municipio/Parte.

Las estimaciones se generan utilizando estimadores de razón que utilizan como variable auxiliar las Estimaciones de la Población incluidas en información de la estadística oficial emitida por el Departamento de Demografía de la Oficina Nacional de Estadísticas.

3.3 Métodos para la asignación de la muestra

Para la **Afijación de la muestra** o lo que es lo mismo, la distribución de la muestra por estratos se realizó una Afijación Proporcional al Tamaño, dado este por el total de población según los estimados actuales de Demografía, debido a que no se cuentan con estadísticas confiables de la cantidad de viviendas del país

Por provincias la distribución de la muestra se realizó con una Afijación de Compromiso entre la Afijación Proporcional al Tamaño, dado este por el total de población y la Afijación Uniforme y el cálculo del tamaño de muestra requerido para los requerimientos de las encuestas periódicas del país, que son, la Encuesta Nacional de Ocupación y la Encuesta de Situación Económica de los Hogares.

3.4 Procedimientos para la selección de la muestra

En Cuba los procedimientos de selección de la Muestra, se definen de acuerdo a los requerimientos particulares de cada encuesta.

Para el proceso de selección de las muestras, las Naciones Unidas recomiendan el Muestreo Probabilístico para las estadísticas oficiales que producen los Órganos Nacionales de Estadísticas.

El principio básico en el que se sustenta el Muestreo Probabilístico, define que cada elemento de la población o unidad de muestreo tiene una probabilidad conocida y No nula de ser seleccionado.

La selección se puede realizar con probabilidades iguales o con probabilidad proporcional al tamaño.

3.5 Diseños en tres etapas: ventajas y desventajas

En Cuba el diseño muestral corresponde a un Muestreo por conglomerados trietápico con estratificación de las Unidades Muestrales de Primera etapa.

DISEÑOS METODOLÓGICOS DE LA ENCUESTA DE HOGARES DE PROPÓSITOS MÚLTIPLES EHPM

DIRECCIÓN GENERAL DE ESTADÍSTICA Y CENSOS (DGEC)
EL SALVADOR

ÍNDICE

Página

Introducción 180
Diseño Muestral EHPM 1998 180
1.- Objetivos de la Encuesta: Generales y Específicos 180
2.- Cobertura de la Muestra (Población Bajo Estudio) 181
3.- Niveles de Estimación 181
4.- Marco Muestral y su Actualización 181
5.- Tipo de Muestreo 182
6.- Estratificación 182
7.- Unidades Muestrales 182
8.- Unidades de Estudio 184
9.- Tamaño de la Muestra de Hogares 184
10.- Tamaño de la Muestra de Segmentos 184
11.- Afijación de la Muestra 185
12.- Proceso de Selección Muestral 185
13.- Empadronamiento (o Enumeración) de Viviendas u Hogares de los Segmentos 186
14.- Estimadores Propuestos 186
14.1 Un antecedente básico: la media muestral por elemento en la unidad primaria 186
14.2 Probabilidad de selección de la unidad primaria (segmento) en el estrato 187
14.3 Fracción de muestreo o probabilidad de selección de una vivienda en el estrato 187
14.4 factor de expansión en el estrato estimador de total 187
14.5 Estimador de la media aritmética 188
14.6 Errores de muestreo por método de conglomerados últimos 188
14.6.1 Estimador de la varianza del total 188
14.6.2 Estimador de la varianza de la media y de la razón 189
14.6.3 Error estándar, coeficiente de variación, error de muestreo e intervalo de confianza 189
15.- Actualización del Marco Muestral 189
16.- Rotación de la Muestra 189
17.- Tratamiento de las No Respuestas 190

18.- Estimadores Ajustados por las Proyecciones de Población por Grupos Sexo-Edad 190
Encuesta de Hogares de Propósitos Múltiples 2003-2007 191
Propuesta de diseño muestral y selección del marco apropiado 191
Objetivos 191
Selección del marco muestral 192
Propuesta de diseño muestral 194
Niveles de desagregación de resultados 197
Sorteo de las unidades primarias de muestreo 197
Sorteo de las unidades de segunda etapa (viviendas) 198
Anexo 1 199
Costos variables unitarios por estrato para la realización de listados cartográficos y de las encuestas de hogares 199
Anexo 2 201
Tamaño de la muestra de la EHPM-2003 por estrato 201

Introducción

Se presenta un resumen de los diseños metodológicos de la Encuesta de Hogares de Propósitos Múltiples EHPM, el primero se refiere al diseño actual de la encuesta EHPM de 1998, elaborado por Froilán Fernández, Profesor del CIENES/OEA, y Daniel Flores De Paz, Muestrista de la División de Información de la DIGESTYC, el segundo se refiere a la propuesta inicial del Diseño Muestral de la EHPM de 2003, presentada por David Glejberman consultor en Muestreo (Uruguayo), con quién he participado como contraparte Nacional.

La Encuesta de Hogares de Propósitos Múltiples comenzó a ejecutarse en El Salvador a partir de 1975, a nivel nacional y con una periodicidad semestral. Posteriormente, el programa de encuestas sufrió algunas modificaciones de cobertura geográfica y de frecuencia en su ejecución, a causa del conflicto vivido en el país, pero en los últimos años su cobertura geográfica nuevamente abarca todo el territorio nacional.

El apoyo financiero y técnico recibido por este programa de encuestas, de parte de organismos internacionales, - en un comienzo el PNUD y el FNUAP, más adelante AID y el Banco Mundial y ahora el BID -, reflejan la gran importancia que éste tiene en el diagnóstico permanente de la situación social del país, constituyéndose en un valioso apoyo, tanto para el Gobierno como para los organismos internacionales, para encauzar la inversión social con eficiencia y efectividad.

En esta nueva etapa de la Encuesta de Hogares de El Salvador, el BID, con el apoyo del BM y CEPAL, ha puesto en marcha el Programa MECOVI "Mejoramiento de las Encuestas de Condiciones de Vida en América Latina y el Caribe", que tiene como objetivos principales fortalecer los Sistemas de Encuestas, para obtener información más completa, en apoyo a la formulación de políticas en el área social y capacitar a los técnicos nacionales en la operación de estas encuestas, en el análisis de la información de las mismas, para el diseño, seguimiento y evaluación de políticas orientadas a reducir la pobreza.

Diseño Muestral EHPM 1998

1. Objetivos de la Encuesta: Generales y Específicos

Objetivos Generales

- Conocer la situación y evolución socioeconómica de los hogares y la población del país.
- Contribuir a la elaboración, seguimiento, y evaluación de las políticas y programas sociales - globales y sectoriales- que componen la inversión y el gasto social.
- Apoyar la descentralización de la gestión pública, en los ámbitos de diagnóstico y toma de decisiones.

Objetivos Específicos

- Realizar diagnósticos respecto de la situación de vivienda, salud, empleo, ingresos y educación, a nivel nacional, regional, departamental, urbano-rural y municipal, e identificar los principales cambios experimentados en cada uno de estos ámbitos, respecto a mediciones anteriores.
- Estudiar la cobertura y nivel de focalización de las principales políticas y programas sociales, especialmente aquellas dirigidas a la población carenciada y grupos prioritarios, estableciendo la demanda insatisfecha de cada uno a nivel nacional, regional y local, así como las tendencias de cambio respecto a mediciones anteriores.
- Realizar diagnósticos relativos a la situación socioeconómica de grupos y sectores prioritarios, estableciendo su evolución y cambios respecto a análisis anteriores.
- Cuantificar la pobreza por el método de las necesidades básicas insatisfechas y establecer los perfiles socioeconómicos de los distintos estratos de hogares.

- Mejorar la medición del ingreso, para preparar estudios sobre la distribución y composición del ingreso del conjunto de los hogares y población del país.
- Disponer de información que permita estimar el impacto de la inversión y del gasto social sobre las condiciones de vida de la población.
- Cuantificar y caracterizar -social y territorialmente- los hogares y la población en situación de indigencia y de pobreza, estableciendo los cambios con respecto a las mediciones realizadas en Encuestas anteriores, e identificar los factores asociados a tales cambios.
- Realizar estudios específicos respecto a cambios socioeconómicos en los hogares y población de los municipios.

2. Cobertura de la Muestra (Población Bajo Estudio)

La población representada en la muestra corresponde a las viviendas particulares y a los hogares y personas que allí habitan, con una cobertura nacional.

3. Niveles de Estimación

Los niveles de estimación geográfico, para los cuales la muestra proporciona información válida, son los siguientes:

Niveles Geográficos	Total	Urbano	Rural
País	X	X	X
Región	X	X	X
Departamento	X	X	X
Municipios Autorrepresentados	X	--	--

4. Marco Muestral y su Actualización

El marco muestral se basa en los antecedentes del último Censo de Población y Vivienda (1992), en cuanto a material cartográfico, como información de población y vivienda.

La división político administrativa de El Salvador contempla 14 Departamentos, los que se subdividen en un total de 262 municipios. Cada municipio, a su vez, se divide en un área urbana (cabecera municipal) y áreas rurales llamadas cantones. Sin embargo, 5 cantones con notorias características urbanas se incluyen en el área urbana del país:

1. Cara Sucia, en Ahuachapán
2. San José La Majada, en Sonsonate
3. Lourdes, en La Libertad
4. San José Las Flores, en San Salvador y
5. Villa Mariona, en San Salvador

Para una mejor cobertura geográfica se divide el país en 5 Regiones, estando éstas constituidas por los Departamentos que a continuación se detallan:

- REGIÓN I (Occidental)
  Santa Ana, Ahuachapán, Sonsonante.
- REGIÓN II (Central I)

La Libertad (*), San Salvador (**), Chalatenango, Cuscatlán.

- REGIÓN III (Central II)
  San Vicente, La Paz, Cabañas.
- REGIÓN IV (Oriental)
  Usulután, San Miguel, Morazán, La Unión.
- ÁREA METROPOLITANA SAN SALVADOR (AMSS)

Constituida por:

Área Urbana de San Salvador y los Municipios de Mejicanos, San Marcos, Ayutuxtepeque, Cuscatancingo, Ciudad Delgado, Ilopango, Soyapango, Antiguo Cuscatlán y Nueva San Salvador.

Se utiliza como marco muestral la cartografía, proporcionada por la misma Dirección General de Estadística y Censos, estableciendo un marco final de 1,073,172 viviendas.

El área urbana de este marco fue actualizado en 1995, por la DIGESTYC, siendo conveniente hacerlo también este año (1997), en las principales ciudades, mediante visitas directas al terreno.

5. Tipo de Muestreo

El tipo de muestreo es estratificado, por conglomerados y polietápico. En cada estrato se obtuvo una muestra independiente que lo representara. El conjunto de todos ellos representa el país.

6. Estratificación

La estratificación utilizada es de tipo geográfico. El país se dividió en 105 estratos (Ver Anexo Uno) y se entendió como tal a la conjunción de división político-administrativa (municipio o agrupación de municipios) y área geográfica (urbana o rural).

Independientemente, los municipios se clasificaron en autorrepresentados y correpresentados.

Municipios Autorrepresentados: Son aquellos que deben estar presentes en la muestra, en atención a la importancia que se les asignó, desde el punto de vista socioeconómico. Además, es de interés tener, para éstos, estimaciones independientes. En total, son 40 los municipios que están en esta categoría y los mismos representan el 64% de las viviendas del país.

Municipios Correpresentados: Son el resto de los municipios que no fueron autorrepresentados. Dentro de un Departamento, se forman dos estratos correspondientes a las áreas urbana y rural de dichos municipios y se seleccionan algunos de ellos en cada estrato con probabilidad proporcional al tamaño, medido en número de viviendas que contiene.

7. Unidades Muestrales

CASO 1:

(Diseño Bietápico) Municipios Autorrepresentados

UNIDADES PRIMARIAS DE MUESTREO (UPM)

Tanto en el estrato urbano como en el rural, las UPM están constituidas por los segmentos de empadronamiento censal (conglomerados de 70 viviendas aproximadas), claramente delimitados.

(*) Se excluyen las áreas urbanas de los municipios de Antiguo Cuscatlán y Nueva San Salvador.
(**) Se excluyen las áreas urbanas de los municipios que conforman el AMSS.

UNIDADES SECUNDARIAS DE MUESTREO (USM) O UNIDADES DE ÚLTIMA ETAPA

ALTERNATIVA A:

Sólo se prepara un listado de direcciones de viviendas particulares en cada UPM

En muchas investigaciones muestrales, no es posible realizar una enumeración de hogares privados en los segmentos seleccionados, dada la escasez de tiempo y/o presupuesto para preparar la muestra en forma exhaustiva, debiéndose conformar el muestrista con preparar un simple listado de direcciones de viviendas particulares, sin haber podido indagar sobre los hogares existentes en cada vivienda. Evidentemente preparar un listado de direcciones de viviendas es mucho más fácil que golpear cada puerta, indagando por los hogares allí existentes.

En esta alternativa, las USM están constituidas, tanto en el área urbana como la rural, por las **viviendas particulares** que existan al momento de la actualización (o preparación de los listados de direcciones). Dentro de estas USM no se realizó submuestreo alguno, ya que se tomó a todos los hogares y a todas las personas que tenían su residencia habitual en las mismas.

ALTERNATIVA B:

Se prepara una enumeración de hogares privados en cada UPM seleccionada

El caso ideal es realizar una enumeración exhaustiva de hogares privados en los segmentos seleccionados (UPM), sobre la base de la cual se efectuará la selección de los hogares a incluirse en la muestra. Esto implica un mayor trabajo, pero si el tiempo y el presupuesto lo permiten habría que emplear este procedimiento.

CASO 2:

Diseño Trietápico
Para los Municipios No Autorrepresentados (Correpresentados).

UNIDADES PRIMARIAS DE MUESTREO (UPM)

En este caso, las UPM están constituidas en el resto del Departamento, una vez descontados los municipios autorrepresentados, por la parte urbana de los municipios correpresentados o por la parte rural de los mismos, según se trate del Estrato Urbano o Rural correspondiente.

UNIDADES SECUNDARIAS DE MUESTREO (USM)

Tanto en el estrato urbano como en el rural, las USM están constituidas por los segmentos de empadronamiento censal (conglomerados de 70 viviendas aproximadas).

UNIDADES TERCIARIAS DE MUESTREO (UTM) o UNIDADES DE ÚLTIMA ETAPA

Las unidades terciarias de muestreo podrán estar constituidas por las viviendas particulares que existan al momento de la actualización de los segmentos o, en el mejor de los casos, por los hogares privados existentes en el segmento muestral correspondiente. (Ver punto anterior ya tratado sobre las dos alternativas para considerar Unidades de Última Etapa, en el caso de los municipios autorrepresentados).

Nota:
El caso 1 (Diseño Bietápico), podría considerarse como Diseño Trietápico, asumiendo que la parte urbana y rural de los municipios autorrepresentados son unidades primarias con probabilidad de selección igual a 1. Esto tiene ventajas para uniformar los procedimientos.

8. Unidades de Estudio

La unidad última de muestreo, tanto para el caso 1 como para el caso 2, anteriormente presentados, permite establecer diversas unidades de estudio. En este caso, las unidades de estudio son las viviendas, hogares y personas.

9. Tamaño de la Muestra de Hogares

En la determinación del tamaño de la muestra se contemplaron los siguientes aspectos:

- Disponer, en algunos municipios autorrepresentados, de un tamaño de muestra que permita dar estimaciones con una confiabilidad mínima aceptable.
- Disponer, en cada región, área urbana y área rural, de un tamaño de muestra que permita dar estimaciones con una confiabilidad bastante adecuada. Evidentemente mucho mayor que la de los municipios muestrales, tomados individualmente.
- En cada estrato debe haber un mínimo de unidades de primera etapa, que permita el cálculo de varianza.
- El costo debe ser aceptable y se debe considerar, además, que en el área rural éste es mayor que el de la parte urbana de la muestra.

Los requerimientos mínimos de precisión, a nivel municipal, son de un error absoluto del 6% y un nivel de confianza del 95%, suponiendo varianza máxima. En todo caso, existen diversas alternativas de tamaño de muestra, de acuerdo a las precisiones exigidas. En este sentido, se preparó un **modelo de simulación** a objeto que las autoridades de la DIGESTYC decidan sobre la alternativa de tamaño muestral, factible de poder llevarse a cabo.

10. Tamaño de la Muestra de Segmentos

La conformación de las Unidades Primarias de Muestreo (en los municipios autorrepresentados) y las Unidades Secundarias de Muestreo (en los correpresentados), debió hacerse sobre la base de los segmentos censales (70 viviendas aproximadas), dado que no se pudo contar con antecedentes censales a más bajo nivel geográfico, como era la intención de los autores.

Por lo tanto, no se pudieron constituir dichas unidades muestrales sobre la base de manzanas (parte o unión de ellas), con un tamaño más para su manejo (25 viviendas aproximadas). Si este hubiese sido el caso, se podría haber seleccionado un mayor número de UMP (o USM), pues la etapa de enumeración se simplifica al tener éstas un menor número de viviendas.

Para determinar el número de unidades últimas de muestreo (viviendas u hogares) a seleccionar dentro de cada UPM o USM según sea el caso, se consideraron aspectos empíricos sobre:

- experiencia anterior,
- intenciones de mensualización de la muestra,
- número múltiplo de bloques de Unidades Últimas,
- homogeneidad y costos,
- facilidades operativas para la realización del trabajo en terreno.

Y es así como se determinó consensualmente, tener -en los segmentos elegidos- un tamaño de 12 viviendas en el área urbana y 16 en el área rural, lo cual significa un determinado tamaño de la muestra de segmentos. Si se hubiese decidido una mayor dispersión de la muestra de hogares, ello implicaría un mayor número de segmentos y un aumento de los costos.

El **modelo de simulación**, ya mencionado, permite utilizar otros parámetros, en cuanto al número de viviendas a seleccionar en los segmentos, lo que podría hacer variar el número de ellos que entraría en la muestra.

11. Afijación de la Muestra

La distribución de la muestra de viviendas, en lo urbano y rural de los municipios autorrepresentados, se hizo con afijación óptima.

$$m_h = m \frac{N_h S_h / \sqrt{c_h}}{\sum (N_h S_h / \sqrt{c_h})}$$

Por ejemplo, la determinación del tamaño de la muestra urbana y de la muestra rural, se determina de la manera siguiente:

$$m_u = m \frac{N_u S_u / \sqrt{c_u}}{N_u S_u / \sqrt{c_u} + N_r S_r / \sqrt{c_r}}$$

$$m_r = m \frac{N_r S_r / \sqrt{c_r}}{N_u S_u / \sqrt{c_u} + N_r S_r / \sqrt{c_r}}$$

Para ello se consideró, de acuerdo a la experiencia de la actual EHPM, una relación de desviaciones estándares y una relación de costos, entre lo urbano y rural, como sigue:

$$\frac{S_u}{S_r} = 1.3 / 1.0 \qquad y \qquad \frac{c_u}{c_r} = 0.5 / 1.0$$

donde:

S_u Desviación estándar (error estándar) urbana
S_r Desviación estándar (error estándar) rural

c_u Costo de una entrevista en la parte urbana
c_r Costo de una entrevista en la parte rural

12. Proceso de Selección Muestral

UNIDAD PRIMARIA DE MUESTREO (UPM)

En los municipios auto representados, las UPM (segmentos en lo urbano y en lo rural) se deben seleccionar dentro de cada estrato, con probabilidad proporcional al tamaño (PPT), medido este tamaño por el número de viviendas particulares obtenidas en el Censo de 1992 ó, en sus casos pertinentes, en el número de viviendas captado en la actualización de 1995 ó 1997.

En la parte no autorrepresentada de cada Departamento, se seleccionan las UPM (Parte urbana o parte rural de cada municipio, según sea el caso) con probabilidad proporcional al tamaño (PPT). Los segmentos dentro de cada UPM también deben seleccionarse con PPT.

UNIDAD DE ÚLTIMA ETAPA (VIVIENDAS U HOGARES)

Las unidades de última etapa, es decir las viviendas (u hogares), se seleccionan aleatoriamente en bloques móviles. El tamaño de la muestra en lo urbano es de 12 viviendas (u hogares) en cada segmento y de 16 en los segmentos rurales, que se seleccionan en bloques agrupando 4 viviendas (u hogares) cada uno de ellos. Es decir, en los segmentos urbanos se eligen 3 bloques y en los rurales 4 bloques.

13. Empadronamiento (o Enumeración) de Viviendas u Hogares de los Segmentos

Los segmentos seleccionados se empadronan antes del levantamiento de la encuesta, para actualizar el número de viviendas y hogares, con la finalidad de detectar los cambios experimentados entre 1992 y el período de la encuesta y determinar, así, las viviendas particulares ocupadas y los hogares correspondientes al momento de la encuesta, las que constituyeron el marco muestral actualizado para seleccionar los bloques móviles de hogares.

Este número actualizado de viviendas (u hogares), debe emplearse en el cálculo de probabilidades y, por ende, en el cálculo de los factores de expansión de la muestra.

$$M'_{hi}$$

Por otra parte, cabe mencionar que las viviendas (u hogares) a seleccionar en cada UPM

$$m_{hi}$$

(doce en el área urbana y dieciséis en el área rural) se seleccionan respectivamente en 3 y 4 bloques móviles. Para ello se calcula un período de selección (K) para 3 ó 4 puntos, según sea el caso. Dentro de ese Período "K", se selecciona un arranque aleatorio, al cual se le suma 1,2 y 3 veces el Período "K", para obtener el punto inicial de los bloque móviles de 4 viviendas (u hogares). Cada bloquecito está conformado por una serie compacta de 4 viviendas (u hogares) sin solución de continuidad, denominada "bloque móvil". Al tomarse esta decisión se pesó cuidadosamente los pro y los contra que esto involucraba. (Los procedimientos comparados fueron: sistemático, aleatorio, bloque fijo y bloque móvil). (Recordar Proceso de Selección Muestral, Unidad de Última Etapa: Viviendas).

14. Estimadores Propuestos

14.1 Un antecedente básico:

La media muestral por elemento en la unidad primaria

(Ejemplo: Ingreso medio estimado por vivienda)

De un modo general, los estimadores propuestos se han construido como algoritmos que deben ser alimentados con las medias muestrales que entreguen las unidades primarias seleccionadas.

En esta sección, se describe el modelo bietápico, entendiéndose que el trietápico es una extensión natural del primero, el cual, aparte de ser más simple, tiene mayor aplicación aquí, en conformidad con los objetivos del estudio.

Para hacer más concreta la exposición, se ha tomado como ejemplo el cálculo del ingreso medio estimado por vivienda, dentro de la Unidad Primaria de Muestreo (UPM).

$$\bar{y}_{hi} = \frac{\sum_{j=1}^{m_{hi}} \sum_{k=1}^{G_{hij}} y_{hijk}}{m_{hi}}$$

donde:

m_{hi} : N° de viviendas logradas en la UPM i del estrato h.

G_{hij} : N° de hogares logrados en la vivienda j de la UPM i del estrato h.

y_{hijk} : Ingreso mensual en el hogar k de la vivienda j de la UPM i del estrato h.

14.2 Probabilidad de selección de la unidad primaria (segmento) en el estrato.

La selección de la UPM se hizo con probabilidad proporcional a su tamaño, medido en número de viviendas particulares ocupadas con moradores presentes a la fecha del Censo de Población y Vivienda de 1992

$$n_h \cdot \frac{M_{hi}}{M_h}$$

donde:

n_h : Número de UPM seleccionadas en el estrato h.

M_{hi} : Número de viviendas en la UPM i del estrato h, según Censo de 1992 (o su actualización en 1995 ó 1997)

M_h : Número de viviendas en el estrato h.

14.3 Fracción de muestreo o probabilidad de selección de una vivienda en el estrato.

La fracción de muestreo o probabilidad de selección de una vivienda es igual a

$$f_{hi} = \frac{n_h M_{hi}}{M_h} \frac{m_{hi}}{M'_{hi}}$$

donde:

M'_{hi} : Número de viviendas empadronadas en la fecha de enumeración de la UPM i del estrato h.

14.4 Factor de expansión en el estrato. estimador de total

El factor de expansión es igual al valor recíproco de la fracción de muestreo, es decir

$$F_{hi} = \frac{1}{f_{hi}} = \frac{M_h}{n_h M_{hi}} \frac{M'_{hi}}{m_{hi}}$$

El total muestral de la UPM i se define como

$$y_{hi} = m_{hi}\,\bar{y}_{hi}$$

y el estimador de total del estrato h

$$\hat{Y}_h = \sum_{i=1}^{n_h} F_{hi}\,y_{hi} = \sum_{i=1}^{n_h} \frac{M_h}{n_h M_{hi}} \frac{M'_{hi}}{m_{hi}}\,y_{hi} = \frac{1}{n_h}\sum_{i=1}^{n_h} M_h \frac{M'_{hi}}{M_{hi}}\,\bar{y}_{hi}$$

$$= \frac{1}{n_h}\sum_{i=1}^{n_h} \hat{M}_{h(i)}\,\bar{y}_{hi} = \frac{1}{n_h}\sum_{i=1}^{n_h} \hat{Y}_{h(i)}$$

donde:

$\hat{M}_{h(i)}$: Número estimado de viviendas en el estrato h, a partir de la actualización de la UPM i.

$\hat{Y}_{h(i)}$: Total estimado en el estrato h, a partir de la información recogida en la UPM i.

14.5 Estimador de la media aritmética

(Ejemplo: Ingreso medio mensual por hogar)

De manera análoga a la que se empleó para estimar el ingreso total en el estrato h, se procede para estimar el número total de hogares existentes en él.

Si este número resulta ser

$$\hat{X}_h$$

el ingreso medio mensual estimado por hogar en dicho estrato será

$$\hat{\bar{Y}}_h = \frac{\hat{Y}_h}{\hat{X}_h}$$

Esta forma de calcular la media aritmética obedece al hecho de que tanto el ingreso mensual por hogar como el número de hogares por vivienda son, ambas, variables aleatorias, de tal manera que presenta la estructura de un estimador de razón.

14.6 Errores de muestreo por método de conglomerados últimos

14.6.1 Estimador de la varianza del total

$$v(\hat{Y}_h) = \frac{1}{n_h\,(n_h-1)} \sum_{i=1}^{n_h} (\hat{Y}_{h(i)} - \hat{Y}_h)^2$$

Esta fórmula es equivalente a la fórmula (4) del Método PC CARP, que aparece en la Pág. 98, Anexo F, de la Guía de Usuario del CENVAR.

14.6.2 Estimador de la varianza de la media y de la razón

$$v(\hat{R}_h) = \frac{1}{n_h\,(n_h-1)\,\hat{X}_h^2} \sum_{i=1}^{n_h} (\hat{Y}_{h(i)} - \hat{R}_h\,\hat{X}_{h(i)})^2$$

Esta es la fórmula (11.39) del Cochrane, en su edición 1978 CECSA. Además esta fórmula es equivalente a la fórmula (5) del PC CARP, Pág. 99, Appendix F, CENVAR User's Guide.

14.6.3 Error Estándar, Coeficiente de Variación, Error de Muestreo e Intervalo de Confianza

Ya con la varianza calculada para el estimador respectivo, se procede a calcular el error estándar (como raíz cuadrada de la varianza); el coeficiente de variación (cuociente entre la desviación estándar y el estimador calculado); el error de muestreo con determinado nivel de confianza (z veces la desviación estándar) y el intervalo de confianza (el estimador calculado, más menos, el error de muestreo). Todo lo cual permite analizar la confiabilidad estadística de las estimaciones provenientes de la EHPM.

La DIS ya viene procediendo con estos cálculos, utilizando el software PC CARP, presentando diversas tablas para analizar la confiabilidad de las cifras, para distintos niveles geográficos, ante distintas alternativas de valores poblacionales (estimador de totales de clase) y porcentajes.

15. Actualización del Marco Muestral

Cuidar la calidad de la encuesta exige, no sólo el diseño de una buena muestra probabilística, sino que además se debe cuidar la selección y adecuado adiestramiento del personal de campo, diseñar cuestionarios cómodos de diligenciar, evitar duplicaciones de preguntas, promover la colaboración de la población para responder la encuesta y mantener una actualización periódica del marco muestral que permita agregar en la muestra de hogares, los nuevos sectores de población que hayan surgido.

Esta actualización del marco debe hacerse cada 2 años (ahora correspondería 1999), de acuerdo al procedimiento que está siguiendo la DIS, con la colaboración de la División de Cartografía ([1]).

16. Rotación de la Muestra

En los sistemas de encuestas continuas se realiza una rotación de la muestra de un período a otro de 1/6 ó 1/5, con el objeto fundamental de no cansar a las familias, las cuales de esta manera permanecerían en la muestra un máximo de 6 ó 5 períodos de encuesta.

Este procedimiento es muy justificable, en las encuestas que tienen una periodicidad pequeña, encuestas trimestrales o cuatrimestrales, en las que el hogar es visitado 4 y 3 veces al año, respectivamente. Por lo tanto en la EHPM, no se justifica utilizar este procedimiento, mientras ésta se lleve a cabo anualmente. En un futuro es posible que la periodicidad de la encuesta aumente al semestre o al cuatrimestre y, en ese caso, recién se debería establecer los procedimientos de rotación que son muy fáciles de aplicar.

Además en la actual EHPM de El Salvador, se contempla una actualización del marco (ver punto 15), para incluir cada dos años una muestra de segmentos nuevos que hayan surgido, lo cual "refresca" la representatividad de la muestra, pues en esos nuevos sectores la composición del hogar y sus

([1]) La División de Cartografía tiene a cargo la actualización del marco muestral 1997, recorriendo la parte urbana de los municipios, para incluir los nuevos barrios y urbanizaciones surgidos. Si se continúa este procedimiento cada dos años, estos antecedentes servirán no sólo a los propósitos de la EHPM, sino que facilitará la preparación del próximo Censo de Población del 2000.

características socioeconómicas son especiales (Parejas jóvenes, con niños pequeños, de un estrato que les permite pagar la mensualidad de la vivienda, etc.).

17. Tratamiento de las No Respuestas

Las causas de no respuesta son de tres tipos:

Tipo A: Viviendas ocupadas, en las que no se obtuvo respuesta, ya sea porque los residentes del hogar se rehusaron a contestar, ausencia temporal de los moradores, o porque no se encontró a nadie en el momento de la entrevista

Este tipo de no respuesta no es deseable, pues afecta las estimaciones de las diferentes variables, por lo que deben mantenerse en una magnitud lo más baja posible. Cuando se produzcan estas no respuestas se procurará recuperarlas, mediante comunicaciones especiales y reentrevista a cargo de entrevistadores y supervisores persuasivos.

Tipo B: Viviendas desocupadas (vacante regular; vacante inadecuada; vacante construyéndose; negocio temporal).

Tipo C: Viviendas que ya no existen (vivienda demolida; casa, tienda o trailer mudado y negocio permanente).

Las no respuestas de los tipos B y C se deben a problemas con el marco muestral (defectos en la enumeración y/o falta de actualización).

18. Estimadores Ajustados Por Las Proyecciones De Población Por Grupos Sexo-Edad

Entre los especialistas de encuestas de hogares existen distintas posiciones sobre la conveniencia de ajustar las cifras de la Encuesta de Hogares por datos de población externos, como serían las proyecciones de población.

La mayoría de los investigadores sostiene que es conveniente efectuar dichos ajustes, dado que las respuestas sobre edad, captadas en censos o encuestas, pueden no ser tan exactas, dada la tendencia de muchas personas que al declarar esta variable, la redondean hacia los años cero (0) y los años cinco (5). Para subsanar esta situación existen procesos de ajuste demográfico, que aplican los especialistas en ese campo. Además señalan que hay que considerar la existencia de proyecciones oficiales de población, por lo que las respectivas cifras de la EHPM, no debieran diferir de ellas.

Otros investigadores sostienen, en cambio, que las proyecciones demográficas no son muy rigurosas ni exactas, para poder utilizarlas como referencia válida de las encuestas, sobre todo si se refieren a niveles geográficos desagregados. Esos técnicos piensan que la encuesta de hogares arrojaría mejores cifras de estructura poblacional por sexo-edad que las proyecciones demográficas, sobre todo si se cuenta con una muestra representativa. Sostienen, igualmente, que dichas variables son fáciles de investigar y sugieren, asimismo, que se debe indagar sobre la fecha de nacimiento de los entrevistados para obtener la edad, lo cual evitaría el redondeo de edad hacia los años cero y años cinco. Además, reafirman que si se contara con procedimientos de actualización permanentes del marco muestral (anuales o bianuales) se lograría obtener muy buenas estimaciones muestrales, que no haría necesario realizar dichos ajustes.

Este tema de discusión es muy importante, dado que muchas variables están relacionadas con la condición de sexo-edad de las personas (La tasa de actividad, la tasa de desempleo y otros indicadores, son muy distintos para hombres y mujeres y para los jóvenes y adultos, por ejemplo), y si la muestra de hogares no refleja bien la estructura etaria y de sexo de la población, se obtendrían estimaciones muestrales erróneas.

Si la proyección de población no es muy adecuada, es posible que los demógrafos puedan nutrirse de las estimaciones poblacionales de la encuesta para mejorar sus proyecciones.

Por ahora, en todo caso, se recomienda adoptar el método del ajuste por datos externos a la encuesta, como son las proyecciones de población, asegurándose así que las estimaciones muestrales concuerden con estas cifras oficiales.

En este sentido, los factores de expansión originales deben ser ajustados por los coeficientes respectivos, que se calculan mediante el cuociente entre la Población Proyectada para el punto central del período de investigación y la Estimación de Población obtenida por la fórmula de los estimadores basada en los factores de expansión originales de la muestra.

$$\text{Coef. Ajuste} = \frac{\text{Poblacion Proyectada}}{\text{Poblacion Estimada EHPM}}$$

Estos ajustes para cada grupo sexo-edad se deberían efectuar a nivel de cada **estrato "h" (Urbano y rural de cada municipio autorrepresentado y del conjunto de los municipios correpresentados en cada departamento)**, pero aún no se cuenta con proyecciones oficiales por sexo-edad a ese nivel. Si dicha información no estuviese disponible, podrían emplearse las proyecciones de población departamentales (urbano y rural, por grupos sexo-edad), de la División de Demografía de DIGESTYC.

Encuesta de Hogares de Propósitos Múltiples 2003-2007

Propuesta de diseño muestral y selección del marco apropiado

Objetivos

a) Mejorar la calidad, cobertura y relevancia para fines sociales de la información proveniente de las encuestas de Hogares de Propósitos Múltiples de El Salvador,

b) Mejorar la capacidad institucional en El Salvador en el manejo y el análisis de las encuestas de hogares y medición de las condiciones de vida,

c) Mejorar el diseño de los instrumentos de investigación, particularmente en lo correspondiente al diseño de la muestra,

d) Mejorar la eficiencia de los métodos de muestreo aplicados en la Encuesta de Hogares de Propósitos Múltiples.

Las tareas específicas a desarrollar son las siguientes:

1. Seleccionar el marco muestral apropiado y realizar las recomendaciones para su actualización, a partir de la revisión de la cartografía y las bases de datos disponibles.

2. Proponer un diseño muestral eficiente en un doble sentido: que proporcione estimadores con la mayor precisión posible (eficiencia en sentido estadístico) y que minimice el costo de la investigación estadística (eficiencia en términos de costos).

3. Calcular los tamaños de muestra, por estratos y por subperíodos de tiempo, de acuerdo con los niveles de desagregación y la precisión deseadas por las Autoridades de la Institución.

4. Realizar recomendaciones para la definición de las unidades de primera etapa en la cartografía censal.

5. Realizar el sorteo de las unidades de primera etapa.

En las secciones siguientes se comentan las tareas realizadas y los resultados obtenidos en ocasión de los trabajos realizados por el consultor.

Selección del marco muestral

En virtud del alcance de la EHPM (cobertura nacional) y de los niveles de desagregación deseados (departamental y principales municipios), el único marco muestral apropiado para la investigación es la cartografía censal con la base de datos asociada, la cual contiene información sobre el total de viviendas hasta el nivel del "segmento censal". El segmento censal es una unidad geográfica bien delimitada y reconocible en la cartografía censal, que a la fecha del último Censo de Población contaba aproximadamente con 75 viviendas en el área urbana y 100 viviendas en el área rural.

Estos segmentos censales han tenido una variación importante en el número de viviendas en virtud de varias circunstancias a saber: crecimiento poblacional y construcción de nuevas viviendas, migración interna, destrucción de viviendas como consecuencia de fenómenos naturales tales como huracanes y terremotos. Una parte de estas modificaciones en el número de viviendas por segmento ha sido captada a través de los trabajos de actualización que el Departamento de Cartografía realiza con cierta frecuencia para las propias Encuestas de Hogares y también como consecuencia de la toma de conocimiento de la construcción de un número importante de viviendas en algunos municipios. También se ha trabajado para reclasificar algunos cantones como áreas urbanas como consecuencia del crecimiento en la periferia de algunas ciudades importantes. Sin embargo, estos crecimientos no son proporcionales a la población existente a la fecha del último Censo, y tampoco se dispone de información completa de estos cambios, razón por la cual la base de datos asociada a la cartografía censal se encuentra en gran parte desactualizada.

De acuerdo con la información proveniente de los segmentos que se han actualizado en campo (principalmente en 1997 y 2001) se encuentra que algunos de ellos han tenido cambios muy importantes, con crecimientos que llevan el número de viviendas a más de 200, y con destrucción de viviendas que dejaron a algunos segmentos con muy pocas unidades habitacionales. Se deduce de lo anterior que los segmentos censales, tal como están definidos -excepto por los que fueron actualizados y re-segmentados en los últimos años-, presentan una importante variabilidad en su tamaño, lo cual es una desventaja si se los pretende utilizar como unidades de primera etapa en un diseño muestral de dos o más etapas. La principal desventaja tiene relación con la asignación de las probabilidades de selección: sería deseable que todas las viviendas -de un mismo estrato- tuvieran la misma probabilidad de selección. En segundo lugar, desde el punto de vista práctico, se el diseño muestral define un cierto número de viviendas a seleccionar en la última etapa (aquí se proponen 12 viviendas), entonces es necesario asegurar un tamaño mínimo de las unidades de primera etapa que garantice que efectivamente se podrá seleccionar aquel número de viviendas. Téngase presente que el marco de viviendas de la cartografía censal contiene información del total de viviendas, algunas de ellas desocupadas, de manera que si se pretende obtener información de 12 de ellas por unidad primaria, entonces será necesario que dicha unidad contenga un número superior.

Otro de los riesgos proviene del hecho que la cartografía no está actualizada, y aunque lo razonable es esperar que donde había un cierto número de viviendas en 1992 (o en 1997 y 2001, fechas de las últimas actualizaciones importantes) ahora exista un número superior, esto no es necesariamente cierto donde han ocurrido fenómenos naturales de alta intensidad.

En consecuencia, la propuesta de marco muestral a utilizar en la EHPM parte de la cartografía censal, parte de los segmentos censales con información sobre el número de viviendas, esto es, el número de viviendas resultante de la actualización del año 2001 o bien de la estimación de razón que resulta de aplicar, en cada departamento, al número de viviendas de 1997 la variación operada entre 1997 y 2001,y

finalmente la variación esperada entre 2001 y 2003. Con el número estimado de viviendas por segmento al año 2003, se define la unidad primaria de muestreo (UPM) de la siguiente manera.

- La UPM coincide con el segmento censal, siempre que su número de viviendas actualizado a 2003 se encuentre comprendido entre 50 y 200, y lo mismo ocurra con los segmentos censales "vecinos" (próximos desde el punto de vista geográfico).
- Si un segmento censal excede de las 200 viviendas, entonces se divide en tantas UPM, aproximadamente del mismo tamaño, de forma que ninguna de ellas contenga más de 200 viviendas.
- Si un segmento censal contiene menos de 50 viviendas, entonces se conglomera con el o los contiguos (del mismo cantón o municipio) para formar una UPM con un mínimo de 50 viviendas (y que no exceda de las 200).

La definición propuesta para las UPM implica una tarea de "acondicionamiento" del marco, la cual fue realizada íntegramente por el consultor, con la participación del Técnico en Muestreo Daniel Flores De Paz, principalmente para discutir el mejor procedimiento a seguir para realizar la tarea, y solicitar información adicional al Departamento de Cartografía para resolver los casos "raros": segmentos censales con 0 viviendas o con más de 500 viviendas en la base de datos.

En la definición del marco muestral es necesario incorporar la dimensión "estrato", si es que el diseño muestral ha de contemplar una estratificación previa del universo a investigar. Tal es el caso de la EHPM.

La encuesta EHPM 2003 deberá proporcionar información confiable a nivel nacional por área, a nivel departamental y al nivel municipal para los municipios más importantes. La importancia de un municipio se ha medido a través de su tamaño poblacional (proyección al año 2003), de forma que se incluyen en el diseño, como dominios de estudio, todos los municipios con más de 30.000 habitantes en el año 2003. Para dar una respuesta adecuada a los niveles de desagregación deseados, el tamaño de la muestra debe calcularse de forma que garantice una buena precisión a nivel municipal para los municipios auto-representados, y una buena precisión para el área urbana y el área rural de cada departamento. Si se tienen buenas precisiones para estos niveles de desagregación, ello garantiza una mejor precisión a nivel del departamento y a nivel nacional. Para lograr estos objetivos, la forma más eficiente es definir estratos a priori que contemplen los mayores niveles de desagregación para la presentación de resultados. Como el área y el municipio están contenidos en la base de datos de la cartografía, entonces estas características deben ser incorporadas al marco acondicionado. Con estas aclaraciones, los estratos que se definen en el marco tienen las siguientes características:

a) Para cada municipio auto-representado se tienen dos estratos: uno formado por las UPM urbanas, y otro por las UPM rurales del municipio.

b) Si un departamento tiene K municipios auto-representados, entonces en ese departamento se definen 2.K+2 estratos, 2.K correspondientes a los estratos urbanos y rurales de los auto-representados, y dos estratos más, uno urbano y otro rural, correspondientes a los municipios restantes del departamento (muncipios co-representados).

En el caso de El Salvador se encontró que al año 2003 serán 50 los municipios con más de 30.000 habitantes, los cuales originan teóricamente 100 estratos. A ellos deben agregarse los 28 estratos correspondientes a los municipios co-representados urbanos y rurales. Decimos "teóricamente 100 estratos" porque en la realidad su número es apenas menor en virtud de la inexistencia de segmentos rurales en algunos municipios de La Libertad y San Salvador.

A partir de esta conformación de los estratos resultan dos parámetros relevantes para el cálculo del tamaño de la muestra en cada uno de ellos: el total de UPM y el total de viviendas por estrato.

El marco, acondicionado de la forma en que se ha descrito, una vez calculados los respectivos tamaños de muestra por estrato, se utiliza inmediatamente para la selección de la muestra de unidades de primera etapa, de acuerdo con el diseño que se detalla en la sección siguiente. Pero además, el marco debe conservarse para ser utilizado, conjuntamente con los resultados de la actualización cartográfica y los resultados de la Encuesta propiamente dicha, para determinar los coeficientes de expansión (el mismo para todas las viviendas de una misma unidad de primera etapa), tema que será abordado en el segundo informe del consultor.

En la definición de los estratos urbanos y rurales se tomó en cuenta los cambios que se adoptaron en la Encuestas de Hogares anteriores en relación con la cartografía censal. Sin embargo, las consultas realizadas con el Jefe del Departamento de Cartografía permitieron concluir que no todas las modificaciones incluidas en las Encuestas anteriores eran de recibo, por cuanto en algunos de los segmentos censales continuaban predominando las características rurales. A continuación se presenta un esquema de los criterios adoptados en las encuestas correspondientes a diferentes períodos temporales.

Segmentos censales de algunos cantones considerados rurales en la cartografía de 1992 según tratamiento otorgado en las diferentes encuestas según período temporal

DEPARTAMENTO	CANTÓN	EHPM-2001	EHPM-2003
Santa Ana	Comecayo	URBANO	URBANO
Santa Ana	Cantarrana	URBANO	URBANO
Santa Ana	Primavera	URBANO	URBANO
Santa Ana	El Portezuelo	URBANO	URBANO
Santa Ana	Loma Alta	URBANO	URBANO
Santa Ana	Natividad	URBANO	RURAL
Santa Ana	Ranchador	URBANO	URBANO
Ahuachapán	Cara Sucia	URBANO	URBANO
Sonsonate	San José La Majada	URBANO	URBANO
La Libertad	Lourdes	URBANO	URBANO
San Salvador	San José Las Flores	URBANO	URBANO
San Salvador	San Luis Mariona	URBANO	URBANO
San Miguel	El Jalacatal	URBANO	URBANO
San Miguel	El Papalón	URBANO	URBANO
San Miguel	El Jute	URBANO	URBANO
San Miguel	El Sitio	URBANO	URBANO
San Miguel	Las Lomitas	URBANO	RURAL
San Miguel	El Zamorán	URBANO	URBANO

Para finalizar con esta sección se recomienda al Departamento de Cartografía que elabore un plan de trabajo para incorporar a la mayor brevedad a la cartografía oficial los resultados de la re-segmentación que habrán de realizarse para la definición de las UPM tal como aquí se las define. Dicho plan deberá abordar, de manera prioritaria, la actualización de todos los segmentos censales que en la base de datos figuran actualmente con cero viviendas.

Propuesta de diseño muestral

El diseño muestral que se propone, responde a los objetivos ya enunciados, a los cuales deben agregarse dos cuestiones adicionales. Por un lado, se pretende replicar el diseño por varios períodos anuales (2003-2007) aunque es razonable que en dicho período se lleve adelante un Censo de Población que favorezca un rediseño más eficiente y sujeto a menos errores de cobertura. Por otro lado, se requiere que la muestra se asigne temporalmente de forma que permita obtener estimaciones de los principales parámetros de la investigación con una periodicidad intra-anual. Una vez explicitado el diseño que aquí se propone, se abordarán ambas cuestiones en esta misma sección.

Las características del diseño muestral propuesto son:

- muestreo probabilístico
- estratificación a priori, donde los estratos son dos por municipio auto-representado (formados uno por todas las UPM urbanas y el otro por todas las UPM rurales del municipio) y dos por los municipios no auto-representados (también uno urbano y otro rural)
- selección de las unidades investigadas (hogares, personas) en dos etapas
- en la primera se seleccionan las UPM ya definidas (segmentos censales, segmentos partidos cuando son muy grandes o segmentos conglomerados cuando uno al menos es muy pequeño en términos de viviendas) con probabilidad aproximadamente proporcional a su número de viviendas (actualizado a 2003)
- en la segunda etapa se seleccionan bloques de cuatro viviendas ocupadas (a la fecha de la actualización del listado cartográfico) cada uno mediante muestreo sistemático
- la muestra final está formada por todos los hogares y todas las personas que residen en las viviendas seleccionadas en la segunda etapa.

Las modificaciones del diseño que aquí se propone, respecto de la EHPM-2001, son las siguientes:

- ahora los municipios auto-representados son 50 (en la EHPM-2001 eran 40)
- las unidades de primera etapa, los segmentos (enteros, partidos o conglomerados), se eligen con reposición (esto es relevante, principalmente, en los municipios pequeños, en los que un mismo segmento podrá ser elegido dos o más veces en la muestra)
- el tamaño de la muestra en la segunda etapa será de 12 viviendas (3 bloques de 4 viviendas) tanto en áreas urbanas como en áreas rurales (mientras que en las encuestas de años anteriores se realizaron 12 encuestas en áreas urbanas y 16 en áreas rurales).

Estas modificaciones responden tanto a cuestiones prácticas como a consideraciones teóricas. La selección con reposición de unidades primarias se realiza con un doble propósito: para contemplar las características del paquete CENVAR (para estimar la variabilidad de los estimadores de los parámetros) y para tener en cuenta las situaciones generadas con los estratos pequeños, con escasos segmentos y donde el tamaño de muestra podría requerir de un número de unidades primarias superior a la cantidad de segmentos en cada estrato. El tamaño de 12 viviendas en los segmentos rurales (en lugar de 16) responde al hecho que es en las zonas rurales donde más incide el efecto de diseño, por lo que resulta más eficiente aumentar el número de unidades primarias de la muestra reduciendo concomitantemente el número de viviendas en la segunda etapa.

La determinación del tamaño de muestra se realiza de manera independiente para cada subuniverso. Los subuniversos se conforman con la unión de dos estratos: el estrato urbano y el estrato rural de cada municipio auto-representado y lo mismo para el conjunto de los municipios co-representados de cada departamento. Así se tiene, para la EHPM-2003, un total de 64 subuniversos (50 municipios auto-representados y 14 municipios "resto").

Determinado el tamaño de la muestra para cada subuniverso, la muestra se asigna a cada estrato mediante la "afijación óptima", minimizando el costo de la investigación para una precisión dada. La precisión deseada en el caso de los municipios auto-representados es el 7.5% para la estimación de la proporción de hogares pobres en dichos municipios. Los costos variables de las encuestas de hogares se estiman en el Anexo 1 de este informe (fue necesario actualizar los cálculos que ya se habían realizado por este consultor en su informe de septiembre de 2001 en virtud de algunos cambios en el número de integrantes

de los equipos de encuestadores y de no haberse incluido, en aquella oportunidad, las cargas de la Seguridad Social correspondientes a los recursos humanos).

Una vez conocidos los tamaños de muestra mínimos necesarios en los subuniversos, se corrige (generalmente a la baja) el tamaño de la muestra en los estratos correspondientes a los municipios co-representados, para obtener una precisión del 5% en la estimación del mismo parámetro al nivel departamental.

Las fórmulas aplicadas para la determinación del tamaño de muestra de los subuniversos y de cada uno de sus estratos se exponen a continuación. Las mismas fueron corregidas para contemplar la falta de respuesta de acuerdo con la experiencia ya recogida en encuestas anteriores (10% en general y 15% en algunas ciudades de La Libertad, San Salvador y San Miguel). Sin perjuicio de ello, los cálculos realizados por el consultor se encuentran detallados en una planilla electrónica, la cual se entrega a los Técnicos de la División de Información Social conjuntamente con este informe.

$$m_h.n_h = \frac{(M_h.\sigma_h.\sqrt{m_h} / \sqrt{c_1 + c_2.m_h}).(\sum_h M_h.\sigma_h.\sqrt{c_1 + c_2.m_h} / \sqrt{m_h})}{\left(\sum_h M_h\right)^2 \cdot \left(\frac{\text{Precisión deseada}}{1.96}\right)^2}$$

donde:

h = es un índice que representa el estrato

m_h = Número de viviendas a seleccionar en la segunda etapa en cada UPM del estrato (en nuestro caso m_h = 12 en todos los estratos)

n_h = Número de UPM a seleccionar en la primer etapa

M_h = Número total de viviendas en el estrato h en el marco muestral

σ_h = Desviación estandar del estimador "proporción de hogares pobres" estimada a partir de la EHPM-2001 ampliada, y corregida por el efecto de diseño

$c_1 + c_2.m_h$ = costo variable de encuestar las m_h viviendas de cada UPM en el estrato h (c_1 es el costo variable unitario de un listado cartografico y c_2 es el costo variable unitario de realizar una encuesta)

1.96 = Valor de una variable con distribución normal estándar que acumula 0.975 de probabilidad.

Si se realiza el cociente $(m_h.n_h)/m_h$ se obtiene el número de UPM a seleccionar en cada estrato. Una vez determinados los tamaños mínimos de muestra necesarios por el procedimiento anteriormente descrito, los mismos se ajustaron para que resultaran múltiplos de cuatro. De esta forma, la afijación temporal de la muestra se realiza asignando aleatoriamente un 25% de la muestra de UPM en cada estrato a cada uno de los cuatro trimestres del año.

La EHPM del año 2004 y siguientes tendrá la misma muestra maestra de UPM, pero la selección de la segunda etapa se realizará excluyendo las viviendas seleccionadas en el año anterior (en los años anteriores). Obsérvese que como la gran mayoría de las UPM contienen un número de viviendas superior a 60, entonces la muestra no se agotará sino hasta el quinto año. En el caso de las UPM con menos de 60 viviendas en el marco, la regla anterior operará hasta el cuarto año.

El cálculo de varianzas de los estimadores será abordado en el segundo informe. Sin embargo, nos adelantamos a decir que, como para estimar varianzas se necesitan al menos dos UPM por estrato, en algunos de ellos no será posible realizar dicha estimación en períodos trimestrales porque el tamaño de muestra es tan solo de una UPM por estrato en el trimestre.

En el Anexo 2 se presenta un resumen de la planilla electrónica que calcula el tamaño de muestra por estrato y departamento.

Niveles de desagregación de resultados

De acuerdo con el diseño muestral adoptado y los períodos de referencia temporales para los cuales se pretende difundir información, debe quedar bien claro que se obtendrán buenas precisiones para el mayor nivel de desagregación (municipios auto-representados) solamente cuando se consideren las encuestas de todo un año. En estos casos la precisión absoluta esperada para estimar una proporción de hogares es del orden del 7.5%, y por tanto, se obtendrán resultados muy pobres para subperíodos dentro del año. El siguiente esquema responde a las recomendaciones que a priori se pueden realizar en virtud del diseño muestral y de los tamaños de muestra calculados según el Anexo 2, sin perjuicio que las mismas están sujetas a los tamaños que efectivamente se logren obtener y de la variabilidad de los estimadores en las respectivas muestras.

NIVEL DE DESAGREGACIÓN	MUESTRA REQUERIDA
Nivel nacional-total	Trimestral
Nivel nacional-urbano	Semestral
Nivel nacional-rural	Semestral
Departamento	Semestral
Departamento-áreas urbanas	Anual
Departamento-áreas rurales	Anual
Municipio auto-representado	Anual

Sorteo de las unidades primarias de muestreo

Conformado el marco muestral tal como se explicó en la sección correspondiente, se trabajó en realizar el sorteo de la muestra en un departamento (Ahuachapán), para dar respuesta a la casuística que podría plantearse en el caso de otros departamentos.

La selección de las UPM en cada estrato se realiza con reposición y con probabilidad proporcional al número de viviendas que cada una de ellas tiene en el marco muestral. Para facilitar la identificación de las unidades sorteadas, en el marco muestral se agregó una columna que cuenta el número de viviendas acumuladas, siguiendo el orden en que las UPM se encuentran en el marco. Se seleccionan mediante una función aleatoria tantos números aleatorios entre 0 y 1 como UPM en la muestra del estrato (según se calcularon en el Anexo 2), luego se multiplica cada número aleatorio por el número de viviendas del estrato, y los resultados obtenidos permiten identificar las UPM de la muestra por simple observación en la columna que acumula viviendas. Este procedimiento asegura el cumplimiento de las dos condiciones mencionadas: selección con reposición y con probabilidad proporcional.

Un cuidado especial debe tenerse con aquellas UPM que no coinciden exactamente con los segmentos censales, es decir, con los que resultan de particiones o conglomeración de segmentos censales. Las UPM sorteadas deberán quedar marcadas convenientemente en el marco muestral, especialmente para las tareas posteriores de expansión de resultados.

En el Anexo 3 se presentan los resultados del sorteo realizado en todos los estratos del departamento de Ahuachapán. Se observará que en cada caso se ha sorteado una UPM suplente, para ser utilizada sólo en casos extremos. La Encuesta no tendrá una buena calidad si por ejemplo el número de unidades de primarias titulares no investigadas supera el 1% de la muestra. También se observa, en la última columna del cuadro del Anexo 3 las características "T" y "TT". La primera significa que la UPM ha sido sorteada como titular una sola vez, mientras que la segunda indica que la UPM ha sido sorteada en la muestra por dos veces.

Sorteo de las unidades de segunda etapa (viviendas)

Para el sorteo de las unidades de segunda etapa se procederá como en las Encuestas anteriores, utilizando el mismo formato de los listados de cartografía. Se recomienda en estos realizar un nuevo diseño privilegiando el espacio destinado a la ubicación (dirección) de las viviendas. En segundo lugar, en la realización del listado cartográfico, se incluirá no sólo las viviendas (con la característica de ocupada, desocupada, en construcción, etc.), sino que se registrará también en el listado toda otra construcción que se encuentre en el segmento censal: locales comerciales, industriales, de servicios o agropecuarios, viviendas colectivas (templos religiosos, hospitales, regimientos militares, etc.), ocupados o desocupados.

Para el sorteo de los tres bloques de cuatro viviendas se agruparán las mismas en el orden del listado, de forma que cada grupo contenga cuatro viviendas ocupadas. Entonces cada grupo estará formado por cuatro viviendas ocupadas más todas las otras construcciones que se encuentren comprendidas entre ellas (viviendas particulares desocupadas, de temporada, en construcción, viviendas colectivas, locales con fines no residenciales).

Seleccionados los bloques, los encuestadores se direccionarán no sólo a las cuatro viviendas ocupadas del bloque, sino que deberán investigar en campo la situación de todas las construcciones del bloque, y realizarán las encuestas en todas las viviendas particulares del área delimitada por el bloque. Esto significa que quizá, en el momento de la Encuesta, algunas viviendas que figuraban como desocupadas hayan pasado a ser ocupadas, y en ese caso, el encuestador deberá realizar más de cuatro entrevistas por bloque. Este nuevo procedimiento (nuevo en relación con las EHPM de años anteriores) no es arbitrario sino que responde a la necesidad de calcular los coeficientes de expansión corrigiendo las cifras del marco mediante la información que los propios encuestadores obtendrán en campo en ocasión de la realización de las encuestas.

ANEXO 1: Costos variables unitarios por estrato para la realización de listados cartográficos y de las encuestas de hogares

Costo mensual de equipo de cartografía			
Concepto	Ciudad de San Salvador	Otras áreas urbanas	Áreas rurales
Gastos de Personal			
Sueldo supervisor de cartografía	$ 611.00	$ 611.00	$ 611.00
Viáticos supervisor	$ -	$ 160.00	$ 388.00
Sueldo motorista	$ 420.00	$ 420.00	$ 420.00
Viáticos motorista	$ -	$ 160.00	$ 388.00
Sueldo 3 cartógrafos	$ 1,316.00	$ 1,316.00	$ 1,316.00
Viáticos 3 cartógrafos	$ -	$ 480.00	$ 1,164.00
Sueldo supervisor de calidad (1/8)	$ 70.00	$ 70.00	$ 70.00
Sueldo verificador (1/4)	$ 108.00	$ 108.00	$ 108.00
TOTAL GASTOS DE PERSONAL	$ 2,525.00	$ 3,325.00	$ 4,465.00
Otros gastos			
Kms. Mensuales	1600	3400	3800
Rendimiento Km/galón	25	30	30
Precio por galón	$ 2.00	$ 2.00	$ 2.00
Costo mensual de gasolina	$ 128.00	$ 226.67	$ 253.33
Otros gastos de vehículos	$ 128.00	$ 226.67	$ 253.33
Gastos en parqueo	$ 20.00	$ 20.00	$ -
Otros vehículos de transporte	$ -	$ -	$ 5.00
TOTAL OTROS GASTOS	$ 276.00	$ 473.33	$ 511.67
GASTO TOTAL MENSUAL	$ 2,801.00	$ 3,798.33	$ 4,976.67
No. Mensual de listados	60	60	24
Gastos de papelería por listado	$ 0.50	$ 0.50	$ 0.50
COSTO UNITARIO APROXIMADO	$ 47.18	$ 63.31	$ 207.36

Costo mensual de equipo de encuestadores			
Concepto	Ciudad de San Salvador	Otras areas urbanas	Areas rurales especiales
Gastos de Personal			
Sueldo supervisor	$ 611.00	$ 611.00	$ 611.00
Viáticos supervisor	$ -	$ 160.00	$ 388.00
Sueldo motorista	$ 420.00	$ 420.00	$ 420.00
Viaticos motorista	$ -	$ 160.00	$ 388.00
Sueldo 3 encuestadores	$ 1,482.00	$ 1,482.00	$ 1,482.00
Viáticos 3 encuestadores	$ -	$ 480.00	$ 1,164.00
Sueldo crítico-codificador (1/4)	$ 119.00	$ 119.00	$ 119.00
TOTAL GASTOS DE PERSONAL	$ 2,632.00	$ 3,432.00	$ 4,572.00
Otros gastos			
Costo mensual de gasolina	$ 128.00	$ 226.67	$ 253.33
Otros gastos de vehículos	$ 128.00	$ 226.67	$ 253.33
Gastos en parqueo	$ 20.00	$ 20.00	$ -
Otros vehículos de transporte	$ -	$ -	$ 5.00
TOTAL OTROS GASTOS	$ 276.00	$ 473.34	$ 511.66
GASTO TOTAL MENSUAL	$ 2,908.00	$ 3,905.34	$ 5,083.66
No. Mensual de encuestas	224	224	224
COSTO UNITARIO DE CAMPO	$ 13	$ 17	$ 23
COSTO IMPRESIÓN CUESTIONARIO	$ 2.00	$ 2.00	$ 2.00
COSTO UNITARIO DE ENCUESTA	$ 15	$ 19	$ 25
Costo de un segm (Listado + encuesta)	$ 226.97	$ 296.52	$ 503.70
Costo unitario total	**$ 18.91**	**$ 24.71**	**$ 41.98**
Raiz cuadrada	4.35	4.97	6.48

ANEXO 2: Tamaño de la muestra de la EHPM-2003 por estrato

Departamento	Municipio	TAMAÑO MUESTRA UPM		
		TOTAL	URBANO	RURAL
AHUACHAPAN	**TOTAL**	**92**	**32**	**60**
AHUACHAPAN	AHUACHAPAN	20	8	124
AHUACHAPAN	SAN FRANCISCO MENENDEZ	20	8	12
AHUACHAPAN	JUJUTLA	20	4	16
AHUACHAPAN	ATIQUIZAYA	20	8	12
AHUACHAPAN	Resto	12	4	8
SANTA ANA	**TOTAL**	**116**	**52**	**64**
SANTA ANA	SANTA ANA	20	16	4
SANTA ANA	CHALCHUAPA	20	12	8
SANTA ANA	METAPAN	20	8	12
SANTA ANA	COATEPEQUE	20	4	16
SANTA ANA	CANDELARIA DE LA FRONTERA	24	8	16
SANTA ANA	Resto	12	4	8
SONSONATE	**TOTAL**	**120**	**60**	**60**
SONSONATE	SONSONATE	20	12	8
SONSONATE	IZALCO	20	8	12
SONSONATE	ACAJUTLA	20	8	12
SONSONATE	NAHUIZALCO	24	8	16
SONSONATE	SAN ANTONIO DEL MONTE	24	16	8
SONSONATE	Resto	12	8	4
CHALATENANGO	**TOTAL**	**52**	**24**	**28**
CHALATENANGO	CHALATENANGO	24	12	12
CHALATENANGO	Resto	28	12	16
LA LIBERTAD	**TOTAL**	**176**	**100**	**76**
LA LIBERTAD	NUEVA SAN SALVADOR	24	20	4
LA LIBERTAD	COLON	24	16	8
LA LIBERTAD	OPICO	20	4	16
LA LIBERTAD	QUEZALTEPEQUE	24	16	8
LA LIBERTAD	ANTIGUO CUSCATLAN	24	20	4
LA LIBERTAD	CIUDAD ARCE	24	8	16
LA LIBERTAD	LA LIBERTAD	24	12	12
LA LIBERTAD	Resto	12	4	8
SAN SALVADOR	**TOTAL**	**320**	**228**	**92**
SAN SALVADOR	SAN SALVADOR	20	20	
SAN SALVADOR	SOYAPANGO	20	20	
SAN SALVADOR	MEJICANOS	24	20	4
SAN SALVADOR	APOPA	24	20	4
SAN SALVADOR	DELGADO	20	16	4
SAN SALVADOR	ILOPANGO	24	20	4
SAN SALVADOR	SAN MARTIN	24	16	8

SAN SALVADOR	CUSCATANCINGO	20	20	
SAN SALVADOR	SAN MARCOS	24	20	4
SAN SALVADOR	PANCHIMALCO	20	4	16
SAN SALVADOR	TONACATEPEQUE	20	4	16
SAN SALVADOR	AYUTUXTEPEQUE	24	20	4
SAN SALVADOR	SANTO TOMAS	24	16	8
SAN SALVADOR	NEJAPA	20	4	16
SAN SALVADOR	Resto	12	8	4
CUSCATLAN	**TOTAL**	**48**	**28**	**20**
CUSCATLAN	COJUTEPEQUE	24	20	4
CUSCATLAN	Resto	24	8	16
LA PAZ	**TOTAL**	**96**	**44**	**52**
LA PAZ	ZACATECOLUCA	24	16	8
LA PAZ	SANTIAGO NONUALCO	24	8	16
LA PAZ	SAN LUIS LA HERRADURA	24	8	16
LA PAZ	Resto	24	12	12
CABAÑAS	**TOTAL**	**72**	**32**	**40**
CABAÑAS	ILOBASCO	24	12	12
CABAÑAS	SENSUNTEPEQUE	24	12	12
CABAÑAS	Resto	24	8	16
SAN VICENTE	**TOTAL**	**44**	**24**	**20**
SAN VICENTE	SAN VICENTE	24	16	8
SAN VICENTE	Resto	20	8	12
USULUTAN	**TOTAL**	**72**	**36**	**36**
USULUTAN	USULUTAN	24	16	8
USULUTAN	JIQUILISCO	24	8	16
USULUTAN	Resto	24	12	12
SAN MIGUEL	**TOTAL**	**72**	**36**	**36**
SAN MIGUEL	SAN MIGUEL	24	20	4
SAN MIGUEL	CIUDAD BARRIOS	24	8	16
SAN MIGUEL	Resto	24	8	16
MORAZAN	**TOTAL**	**48**	**24**	**24**
MORAZAN	SAN FRANCISCO (GOTERA)	24	16	8
MORAZAN	Resto	24	8	16
LA UNION	**TOTAL**	**72**	**28**	**44**
LA UNION	CONCHAGUA	24	4	20
LA UNION	LA UNION	24	16	8
LA UNION	Resto	24	8	16
EL SALVADOR	**TOTAL**	**1400**	**748**	**652**

INFORME TÉCNICO DE LAS EXPERIENCIAS EN DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES

INSTITUTO NACIONAL DE ESTADÍSTICA (INE)
GUATEMALA

ÍNDICE

Página

Síntesis de Informe Técnico de las Experiencias en Guatemala 205
Utilización de proyecciones 206
Actualización cartográfica 206

Síntesis de Informe Técnico de las Experiencias en Guatemala

Dada la premura de tiempo disponible, ya que la solicitud de este informe llegó con un margen estrecho de tiempo en relación a la fecha límite de recepción, los renglones que a continuación se presentan constituyen una versión muy sintética de lo realizado en Guatemala, en parte porque estos trabajos los han efectuado distintas personas con especialidades también diferentes (muestristas, cartógrafos, estadísticos, etc) los cuales no pudieron ser ubicados, a fin de obtener de ellos un informe más detallado y específico.

La cartografía censal fue actualizada en los años 1992 y 1993 para el levantamiento de los Censos Nacionales de Población y Habitación de 1994. Cada sector censal en el área urbana es un área geográfica que contiene alrededor de 200 viviendas; cada sector Censal Rural es un área geográfica que contiene alrededor de 150 viviendas.

El marco que sirvió de base para el diseño de la muestra de las encuestas (ENIGFAM, ENCOVI, ENEI), fue la Información de los sectores censales (11,170). Al tener los sectores seleccionados, se procedió a identificarlos cartográficamente, utilizando el archivo base cartográfico de la sectorización censal. La cartografía de los sectores seleccionados fue actualizada a la fecha de dichas encuestas y se elaboró un listado de viviendas, hogares y total de personas por cada sector cartográfico. Se dibujaron y/o corrigieron los originales de la cartografía, para su actualización. El listado de viviendas y hogares sirvió de marco para la selección de las viviendas a encuestar. Se preparó el material para el encuestador, donde se identificaron las viviendas seleccionadas en mapas listados y croquis de cada sector de interés.

Actualmente se está realizando la cartografía correspondiente al IX censo de Población, VI de habitación y IV agropecuario a realizarse en octubre del año 2002 y abril del 2003 respectivamente. Esta nueva cartografía substituirá a la realizada en 1994 la cual se considera obsoleta. Con este nuevo trabajo se espera que para futuras encuestas, esta nueva cartografía, nos permita diseñar muestras autoponderadas y además para disponer de datos mas precisos en cuanto a la cantidad de personas, hogares y comercios se refiere. El crecimiento no uniforme del número de viviendas por sector en el tiempo, no permite aplicar probabilidades proporcionales al tamaño (PPT) lo que obliga a diseñar muestras *no autoponderadas*, pues el tamaño de los sectores en término de viviendas difiere mucho de los registrados en 1994 y su estimación es compleja e inexacta.

Para llevar a cabo la realización de los procedimientos de estratificación del marco de muestreo se han utilizado, para las encuestas ENIGFAM, ENCOVI Y ENEI el grado de educación como variable proxi, tomando como referencia el censo de población realizado en el año de 1994.

Asimismo se clasificó cada hogar en uno de los tres estratos que se definieron los cuales fueron:

1.- estrato alto urbano
2.- estrato medio urbano
3.- estrato bajo urbano
4.- estrato alto rural
5.- estrato medio rural
6.- estrato bajo rural

Las muestras diseñadas para las encuestas de hogares realizadas fueron elaboradas por especialistas nacionales e internacionales en la materia y el tamaño de ellas se hizo por métodos similares a la asignación óptima de Neymann. (Corchan).

Todas las muestras que se han utilizado en las encuestas realizadas Bietápicas, tomando como referencia para la primera etapa de muestreo el marco de los 11,170 sectores en los cuales se dividió la Republica de Guatemala para el Censo de 1994.

En la Encuesta de Empleo e Ingresos (ENEI 2002) se seleccionaron 491 sectores aleatoriamente, en la primera etapa y en la segunda etapa de muestreo se dividieron los sectores en segmentos compactos de seis viviendas cada uno y se seleccionó aleatoriamente un segmento compacto por cada uno de los sectores seleccionados.

En la Encuesta de Condiciones de Vida (ENCOVI 2000) se seleccionaron 745 sectores aleatoriamente en la primera etapa y en la segunda etapa de muestreo se seleccionaron dos segmentos compactos de 6 viviendas en promedio por cada sector urbano , el primero de los dos fue elegido aleatoriamente y el segundo fue elegido sistemáticamente, a una distancia de ***k*** segmentos donde ***k=total de segmentos del sector/2,*** y un segmento compacto de 12 viviendas en promedio por cada sector rural

En la Encuesta de Ingresos y Gastos Familiares (ENIGFAM 1998) se seleccionaron 432 sectores aleatoriamente y en la segunda etapa de muestreo se seleccionaron aleatoriamente muestras constantes de 18 hogares por cada uno de los sectores urbanos previamente sorteados y 15 hogares por cada sector rural.

La experiencia con la que contamos, hasta el momento, en estas tres encuestas es de muestras de dos etapas, en las cuales la muestra de primera etapa ha sido de conglomerados (sectores) y la de segunda etapa de segmentos compactos y hogares tal como se indicó.

En el caso de ENEI es la primera vez que en Guatemala se aplica una muestra de paneles rotativos. El diseño muestral de la ENEI esta configurado para encuestas periódicas con intervalo trimestral y una rotación cíclica de paneles equivalentes a 1/5 de la muestra total, es decir que cada trimestre se incorpora un 20% de sectores nuevos en la muestra de primera etapa, de tal manera que en 5 trimestres se ha renovado el 100% de la muestra.

Utilización de las proyecciones:

Al momento de realizar los cálculos para las publicaciones de los resultados se han utilizado las proyecciones de población a fin de comparar la población expandida con la proyectada y verificar su proximidad. También se han utilizado las proyecciones de población para ajustar o calibrar los factores de expansión o expansores que sirvieron para realizar las estimaciones y su variabilidad.

Actualización cartográfica:

En vista que la información del censo con la que se cuenta a la fecha, es de el año de 1994, se ha implementado el procedimiento de actualización cartográfica en el cual se designa un grupo de cartógrafos los cuales actualizan los sectores seleccionados en la primera etapa de muestreo y actualizan también los listados de estructuras de estos sectores.

En todos los proyectos se ha utilizado la cartografía tradicional, la cual como mencionamos anteriormente es la cartografía del censo de 1994. Los procesos de selección se realizan de manera tradicional al igual que el calculo de los expansores. Guatemala no utiliza por el momento sistemas digitalizados.

DISEÑO MUESTRAL DE LA XXIV ENCUESTA PERMANENTE DE HOGARES

INSTITUTO NACIONAL DE ESTADÍSTICA (INE)
HONDURAS

ÍNDICE

Página

I.- Introducción 209
II.- Definición 210
III.- Objetivos del Marco Muestras 210
IV.- Cobertura del Marco Muestral 211
V.- Unidades del Marco Muestral 212
VI.- Estratificación del Marco Muestral 216
VII.- Flexibilidad y Ventajas del Marco 217
VIII.- Actualización del Marco Muestral 217

I. Introducción

La mayor parte de las investigaciones orientadas a obtener información de los distintos quehaceres de la sociedad, emplean como principal instrumento de recolección de datos las distintas técnicas de muestreo probabilistico que constituyen ya una parte rutinaria de las diferentes actividades que realizan la oficinas nacionales de estadística. Estas técnicas se basan fundamentalmente en listas de unidades de investigación como ser por ejemplo listado de viviendas, aldeas, zonas, manzanas, segmentos, etc.

La palabra Marco, generalmente es empleada para englobar todos los listados de las diferentes etapas de la investigación, de aquí la denominación de "Marco de Muestreo".

Los diferentes investigadores que aplican el muestreo probabilístico en sus investigaciones, parten del hecho de que ya cuentan con el "Marco de Muestreo" respectivo. Por lo común si bien se tiene el Marco siempre es recomendable hacer una evaluación de la situación en la que se encuentra ya que un "Marco de Muestreo" práctico no sólo proporciona una lista de la cual se pueden seleccionar unidades para su inclusión en la muestra, sino que también debe dar suficiente información sobre cuya base se puedan identificar especialmente las unidades sobre el terreno. De no ser así, se pueden producir distorsiones en las probabilidades de selección y en la estructura de la muestra en general que afectan los resultados que se puedan obtener de la población bajo muestreo.

La elección de un "Marco de Muestreo" adecuado para todas las etapas de la selección de la muestra constituye un aspecto crítico del diseño muestral, ya que de la calidad y confiabilidad del Marco va a depender también la calidad y confiabilidad de los datos a obtenerse de una muestra de la población que se estudia.

El Instituto Nacional de Estadística, desde su creación y dentro de sus actividades de obtener información confiable de bajo costo y de alta calidad proveniente de los hogares de los diferentes estratos sociales y económicos de la población, continua desarrollando la serie de rondas programadas en forma continua de la Encuesta Permanente de Hogares de Propósitos Múltiples y a su vez afrontar un conjunto de diferentes encuestas tanto puntuales como no puntuales orientadas a captar diversas situaciones que presentan los hogares. Con el propósito de optimizar recursos tanto en personal como económicos se construye un único "Marco Muestral" que permita la elaboración de los distintos diseños , así como la elección de las diferentes muestras denominándose a este tipo de marcos **Marco Maestro de Muestreo,** que por comodidad abreviaremos como MMM, este tipo de marcos representa un insumo fundamental para el diseño de un Sistema Integrado de Encuestas a Hogares; es en este contexto también que se generan las denominadas **Muestras Maestras,** definidas como toda muestra a partir de la cual pueden seleccionarse submuestras independientes denominadas réplicas o grupos de unidades con algún nivel de traslape que se emplean para diferentes rondas de una encuesta o para proyectos específicos que se deseen implementar.

En las siguientes líneas presentaremos las definiciones y conceptos que se emplearon para la elaboración del "Marco de Muestreo" para la Encuesta Permanente de Hogares de Propósito Múltiples y las recomendaciones necesarias para la completitud del marco en función de los resultados a obtenerse del Censo Nacional de Población y Vivienda que se efectuó del 28 de Julio al 30 de Septiembre del presente año.

Es necesario puntualizar que el presente Marco se construyo con los datos proporcionados por el Censo de Población y Vivienda, datos procedente de su actualización cartográfica, la misma que necesita ser evaluada y otra vez actualizada con la información en primera instancia del relevamiento censal a los hogares.

Un primer problema al que nos enfrentamos fue creación de un código 3 referido como urbano/rural, es decir poblaciones con características urbanas pero que se siguen considerando como áreas rurales y el censo los creo sólo con fines de su operativo de actualización cartográfica.

Un segundo problema es que los listados presentados de la actualización cartográfica no incluyen el número de personas por vivienda información que se obtiene para cualquier actualización cartográfica, mucho más si se trata del censo. Esto implico que no se tiene un listado de cada una de las viviendas actualizadas, ya sea por sector, segmento, etc.

Finalmente en el listado no se identifica la condición de la vivienda, es decir no se identifica si esta ocupada, desocupada o en construcción, simplemente se tiene el número total de edificaciones sin considerar las características que estas tienen.

Sin embargo es información que nos permite construir el presente Marco Muestral en forma provisional hasta contar con toda la información necesaria para obtener el Marco Muestral definitivo.

II. Definición

Existen variadas y diferentes definiciones de lo que es un "Marco de Muestreo" pero todas concuerdan en que define en forma clara y precisa el alcance y limitaciones que tiene en cuanto a cobertura geográfica, poblacional y unidades que tiene.

La definición mayormente empleada es que el Marco Muestral es un listado de todas las unidades de investigación de un estudio en particular, las cuales están perfectamente definidas e identificadas en forma unívoca.

Complementaremos esta definición de manera tal que satisfaga los requerimientos que nosotros necesitamos para la nueva serie de Encuesta Permanente de Hogares de Propósitos Múltiples que se iniciara a partir de la actualización cartográfica del Censo de Población y Vivienda del 2001, el mismo que deberá complementarse con los datos de las características de los hogares provenientes del relevamiento censal.

Definición.- El "Marco de Muestreo" se define como una lista completa, organizada en forma de base de datos que contiene a todos y cada uno de los elementos de la población de interés que participaran en cada una de las fases de selección de la muestra. El Marco también esta formado por todos los mapas y planos a diferentes escalas que nos permiten identificar en forma precisa y clara los limites físicos que tienen las diferentes unidades de selección, considerándose como parte principal de éste los registros y listados en los que se detalla las referencias que nos permitan identificar en forma exacta las viviendas seleccionadas. Donde todos y cada uno de los elementos de los que esta compuesto el Marco tienen una probabilidad conocida y diferente de cero de ser seleccionados de alguna de las muestras que se puedan extraer del mismo.

El Marco no debe tener unidades repetidas u omitidas y además debe actualizarse en forma permanente de manera tal que siempre refleje los cambios estructurales que se van produciendo en la geografía del país. Tiene que detectarse la conformación de nuevas unidades de selección, su distribución física el cambio que se presenta debido a nuevas edificaciones y construcciones nuevas que se van generando debido al crecimiento de la población.

Basándose en la definición dada es que nosotros desarrollamos el Marco Muestral para la Encuesta Permanente de Hogares de Propósitos Múltiples.

III. Objetivos del Marco Muestral

Los principales objetivos que tiene que cumplir el Marco para que este sea de uso permanente son los siguientes:

- Abarcar todas las unidades de investigación que tiene la población de la encuesta y nos permita seleccionar muestras para varias encuestas o diferentes rondas de la Encuesta Permanente de Hogares.

- Identificar y proporcionar acceso a cada una de las unidades de investigación de la población encuestada, donde además tengan probabilidad conocida de selección o se pueda determinar la misma para su selección en la muestra.
- Permitir la identificación de las áreas de mayor crecimiento urbano como también las áreas rurales o departamentos de mayor emigración de su población.
- Determinar en forma precisa y objetiva la cobertura geográfica y poblacional que la encuesta cubrirá.
- Permitir que los resultados sucesivos de la Encuesta de Hogares sean comparables con las del Censos de Población y Vivienda y con otros estudios estadísticos orientados a obtener información de las condiciones de los hogares particulares.
- Ser un Marco computarizado y de fácil manejo por todos los técnicos relacionados con él.

IV. Cobertura del Marco Muestral

La cobertura geográfica del presente Marco Muestral es Nacional, es decir cubre los 18 departamentos del país en sus áreas urbanas y rurales.

En el cuadro 2, se presenta el número de viviendas urbanas y rurales obtenidas por la actualización cartográfica para el CNPV – 2001, manteniéndose en forma separada las dos ciudades más grandes del país que son Distrito Central y San Pedro Sula.

Cuadro # 2
Viviendas Urbanas y Rurales por Área Geográfica

AREA GEOGRAFICA	VIVIENDAS	%
Distrito Central	189,738	13,14
San Pedro Sula	123,283	8,54
Resto Urbano	384,928	26,67
Rural	735,715	50.96
Total Nacional	1,443,664	100,00

En relación a 1988 se puede observar que del 14% de viviendas que estaban ubicadas en el Resto Urbano en la actualidad prácticamente se duplico el número de las mismas, llegando a representar en estos momentos del total (de viviendas actualizadas) el 26,67%. De la misma forma en el área rural disminuyeron en aproximadamente en un 16%. En el censo de 1988 las viviendas ubicadas en área rural representaban un 66.5% del total nacional.

En el cuadro 3, se presenta el total de viviendas ubicadas en las áreas urbanas y rurales del censo de 1988 y la actualización cartográfica para el censo del 2001.

Cuadro # 3
Total de Viviendas Urbanas y Rurales,
Según Censo de 1988 y 2001

AREA GEOGRAFICA	Censo de Población y Vivienda 1988		Actualización Cartográfica Censo de Población y Vivienda - 2001	
	VIVIENDAS	%	VIVIENDAS	%
Distrito Central	101,350	12,3	189,738	13,14
San Pedro Sula	59,597	7,2	123,283	8,54
Resto Urbano	115,880	14,0	384,928	26,67
Rural	549,875	66,5	735,715	50.96
Total Nacional	826,705	100,00	1,443,664	100,00

Para poder determinar e identificar en forma precisa cuales son las áreas rurales que se están convirtiendo en áreas urbanas es preciso y necesario mantener un único concepto de que se define como área urbana y que área rural. En la actualidad se define como área urbana a toda población que tenga una mínima combinación de servicios como ser agua potable, alumbrado público, correo, centros de salud, escuelas, etc[1]. Además tengan un mínimo de 2,000 habitantes. Toda aldea que cumpla con esas condiciones ya es considerada urbana y el resto de las poblaciones que no cumplan con esta combinación mínima son consideradas poblaciones rurales. Definición dada para el censo de 1988 y que en la actualidad se sigue manteniendo.

En el cuadro 4, se puede observar el total de viviendas por departamento según área geográfica, Actualización Cartográfica para el Censo del 2001.

Cuadro # 4
Total de Viviendas
Según Departamento y Área Geográfica

	VIVIENDAS			%	
DEPARTAMENTO	**TOTAL**	**URBANO**	**RURAL**	**URBANO**	**RURAL**
ATLANTIDA	79,371	42,138	37,233	5.95	5.03
COLON	53,051	16,612	36,439	2.38	4.95
COMAYAGUA	76,074	32,996	43,078	4.71	5.71
COPAN	58,551	20,755	37,796	2.97	5.14
CORTES	294,904	199,832	95,072	28.59	12.89
CHOLUTECA	79,987	25,444	54,543	3.65	7.41
EL PARAISO	70,546	20,639	49,907	3.00	6.75
FRANCISCO MORAZAN	273,664	206,465	67,199	29.58	9.13
GRACIAS A DIOS	12,710	2,426	10,284	0.35	1.40
INTIBUCA	31,978	6,874	25,104	0.98	3.41
ISLAS DE LA BAHIA	10,606	3,993	6,613	0.57	0.90
LA PAZ	30,666	8,635	22,031	1.24	2.99
LEMPIRA	45,893	6,530	39,363	0.94	5.35
OCOTEPEQUE	23,254	7,303	15,951	1.05	2.17
OLANCHO	84,468	24,998	59,470	3.58	8.08
SANTA BARBARA	76,398	20,809	55,589	2.98	7.56
VALLE	32,748	11,549	21,199	1.65	2.88
YORO	101,965	41,134	60,831	5.83	8.25
TOTAL	**1,433,664**	**697,949**	**735,715**	**100.00**	**100.00**

V. Unidades del Marco Muestral

De acuerdo al tipo de investigación que deseamos realizar se define las unidades que debe contener el Marco Muestral. El presente Marco esta orientado a investigar a todas las personas que tienen residencia permanente en una ciudad, aldea o caserío (llegamos a las personas a través de la vivienda particular). En este contexto para nosotros las unidades del marco están constituidas por todas las viviendas particulares ocupadas o no al interior de cada uno de los 18 departamentos del país tanto en su área urbana como rural.

[1] Definición del Censo de Población y Vivienda de 1988, la cual no cambio en el CNPV -2001

Definición:

1) **Vivienda Particular**.- Es aquella destinada al alojamiento de uno a tres hogares, grupo de personas o de una sola persona.[2]

2) **Hogar Particular**.- Es aquel formado por una sola persona o grupo de dos o más personas que se asocian para ocupar, en forma total o parcial una vivienda, así como para proveer el presupuesto para satisfacer las necesidades de alimentación o de otra índole de las mismas.[3]

Basándonos en estas definiciones y con la información de la actualización cartográfica del Censo del 2001, es que empezamos a construir y definir las unidades de las que estará compuesto nuestro marco.

Honduras política y administrativamente se divide en 18 departamentos que a su vez se subdividen en Municipios, Aldeas y Caseríos, entonces las unidades del Marco Muestral tendrán las características siguientes:
En el área urbana y para el área rural se tiene la estructura siguiente:

a. Identificación Geográfica Urbana

Departamento
Municipio
Zona
Aldea
Caserío
Sector
Segmento
Barrio
Manzana

b. Identificación Geográfica Rural

Departamento
Municipio
Zona
Aldea
Caserío
Sector
Segmento

c. Características de la Vivienda*

Tipo de Vivienda
Condición de Ocupación de la Vivienda
Hogares por Vivienda

d. Características del Hogar*

Número de Personas por Hogar y Vivienda
Sexo
Edad
Nivel de Educación
Condición de Empleo

[2] y [3] Manual del Encuestador EPHPM septiembre de 2001.
* Información que tiene que complementarse un vez se tengan los resultados del CNPV-2001

Es necesario resaltar que el Marco tiene que completarse con la información en lo referente a las características de la vivienda y hogar con datos del CNPV del 2001, una vez se termine de procesar el mismo.

El Marco Muestral se ha empezado a conformar con la información por sector y segmento censal proveniente de la actualización cartográfica realizada para el relevamiento censal del 2001. Actualmente se tiene en formato Excel la información del número de viviendas que tiene cada uno de los sectores y segmentos censales, los cuales están debidamente codificados permitiendo su identificación uno a uno como pertenecientes a zona urbana o rural, municipio, aldea, caserío y finalmente al departamento.

Adicionalmente se cuenta con los mapas respectivos por departamento, municipio, aldeas y caseríos en los que se puede identificar la distribución de las viviendas tanto a nivel urbano como rural, facilitando de este modo los operativos de trabajo de campo y selección de la muestra.

La característica fundamental en la construcción de las unidades de las que debe estar compuesto el Marco Muestral consiste en que estas tienen que tener un número suficientes de viviendas que nos permita poder llevar a cabo dos o más encuestas a la vez sin que ninguna de ellas se interfiera entre sí.

La actualización cartográfica para el censo 2001, incluye una nueva área de enumeración censal que es la del **Sector Censal**, que esta compuesta de uno hasta nueve segmentos censales que en promedio tienen 70 viviendas, por lo que el sector censal tiene entre 282 a 350 viviendas en promedio y que además están bien delimitadas y son de fácil identificación en el terreno. Una de las ventajas que se tiene con estos sectores es que tanto en el área urbana como rural tiene el mismo tamaño en viviendas.

De acuerdo al objetivo planteado se determinó que las unidades de primera etapa sean los **sectores censales.**

Las Unidades de Primera Etapa para el Marco Muestral lo constituyen los sectores censales, que como ya hemos visto son de fácil ubicación e identificación tanto en los mapas como en el terreno y tienen un número suficiente de viviendas que nos permite soportar diferentes encuestadas orientadas a los hogares en un mismo tiempo.

Para la construcción de las Unidades de Primera Etapa se siguen los siguientes pasos:

i). En la hoja electrónica denominada Datos Precensales, se separa cada uno de los departamentos por área (urbana o rural) creándose hojas electrónicas para cada uno de los departamentos.

ii). Al interior de cada departamento se ordena la información por municipio, aldea, caserío y finalmente por sector y segmento uniendo lo urbano con lo rural en función del orden en el que están listados, es decir del primer municipio, sector y segmento hasta el que ocupa el último lugar.

iii). Hecho este ordenamiento se empieza a acumular el total de viviendas sumando las viviendas de cada uno de los segmentos de los que esta compuesto cada sector.

iv). Manteniendo el mismo orden del punto ii) se empieza la numeración que identificara a cada una de las UPMs. Esta numeración es en orden decreciente es decir se empieza desde el número 1 hasta el número **n,** numeración correlativa que empieza en el área urbana y termina en el último municipio, sector y segmento del área rural. Además nos indica cuantas UPMs tiene el departamento.

En el cuadro 5, se presenta el total de viviendas y UPMs que tienen cada uno de los departamentos.

Cuadro # 5
Total de Viviendas y UPMs
Según Departamento

DEPARTAMENTO	VIVIENDAS	UPMs
ATLANTIDA	79,371	270
COLON	53,051	189
COMAYAGUA	76,074	250
COPAN	58,551	205
CORTES	294,904	1,032
CHOLUTECA	79,987	280
EL PARAISO	70,546	251
FRANCISCO MORAZAN	273,664	970
GRACIAS A DIOS	12,710	43
INTIBUCA	31,978	118
ISLAS DE LA BAHIA	10,606	39
LA PAZ	30,666	121
LEMPIRA	45,893	171
OCOTEPEQUE	23,254	81
OLANCHO	84,468	283
SANTA BARBARA	76,398	265
VALLE	32,748	119
YORO	101,965	359
TOTAL	**1,433,664**	**5,046**

Al interior de cada una de las Unidades Primarias de Muestreo ya constituidas se define agrupaciones más pequeñas constituidas por grupos de cinco viviendas. Normalmente se conocen como segmentos compactos (SC) a estas unidades, nosotros denominaremos Unidades de Segunda Etapa que por comodidad abreviaremos como USE. En esta etapa se requiere contar con un listado exhaustivo de las agrupaciones, en que se detallen las características del sector permitiendo al encuestador la identificación precisa de cada uno de los segmentos compactos seleccionados para la encuesta.

Para su construcción se siguen los siguientes pasos:

i) Sabemos que cada UPM esta compuesta de uno a nueve segmentos y que estos últimos en promedio tienen 70 viviendas. Los segmentos están ordenados de acuerdo a los criterios de composición del sector y estos no siempre tienen un orden correlativo. Por ejemplo hay un sector que esta compuesto por los segmentos 1, 5, 30 y 56 y así como estos sectores hay muchos más.

ii) Por cada uno de estos segmentos se tiene el total de viviendas. Este total de viviendas se divide entre cinco (viviendas que constituye el SC). Esta división nos dará el número de SC, que tendrá cada uno de los segmentos.

iii) Obtenido el número total de SC, que tiene cada UPM, estos se numeran en forma correlativa desde el número 1 hasta el número **k**, que viene a constituir el último SC que tiene la UPM. Esta numeración se la hace respetando el orden en el que los segmentos se encuentran listados en cada una de las UPMs.

iv) En cada uno de los segmentos se crean intervalos de números cerrados a la izquierda y derecha al que hemos denominado como Rango del Segmento Compacto, que nos indica de

que número a que número van los segmentos compacto y a su vez no dice cuantos segmentos compactos tiene cada segmento.

v) En el siguiente ejemplo de la UPM uno del departamento de Atlántida del municipio La Ceiba se puede apreciar la estructura de los SC al interior de las UPMs.:

UPM	ZONA	SECTOR	SEGMENTO	VIVIENDAS	RANGO DEL SEGMENTO COMPACTO
1	01	01	001	67	1 - 13
1	01	01	002	67	14 - 27
1	01	01	003	61	28 - 39
1	01	01	039	75	40 - 54
1	01	01	040	68	55 - 68

VI. Estratificación del Marco Muestral

De manera muy breve podemos decir que la estratificación se refiere a una subdivisión de una población en subconjuntos con características propias. Este proceso se lleva a cabo siempre antes de seleccionar la muestra.

El criterio de estratificación empleado para el presente marco es el criterio geográfico. El país en su conjunto se divide en dos estratos que son el urbano y rural, heterogéneos entre si y homogéneos en su interior, la estratificación por departamento junto con el número de UPMs y total de viviendas que tienen se presenta en el cuadro 6

Cuadro # 6
Total de Viviendas y Unidades Primarias de Muestreo,
Según Departamento y Área Geográfica

	VIVIENDAS			UPMs		
DEPARTAMENTO	**TOTAL**	**URBANO**	**RURAL**	**URBANO**	**RURAL**	**TOTAL**
ATLANTIDA	79,371	42,138	37,233	144	126	270
COLON	53,051	16,612	36,439	61	128	189
COMAYAGUA	76,074	32,996	43,078	119	131	250
COPAN	58,551	20,755	37,796	78	127	205
CORTES	294,904	199,832	95,072	696	336	1,032
CHOLUTECA	79,987	25,444	54,543	93	187	280
EL PARAISO	70,546	20,639	49,907	77	174	251
FRANCISCO MORAZAN	273,664	206,465	67,199	738	232	970
GRACIAS A DIOS	12,710	2,426	10,284	9	34	43
INTIBUCA	31,978	6,874	25,104	31	87	118
ISLAS DE LA BAHIA	10,606	3,993	6,613	14	25	39
LA PAZ	30,666	8,635	22,031	40	81	121
LEMPIRA	45,893	6,530	39,363	36	135	171
OCOTEPEQUE	23,254	7,303	15,951	27	54	81
OLANCHO	84,468	24,998	59,470	91	192	283
SANTA BARBARA	76,398	20,809	55,589	74	191	265
VALLE	32,748	11,549	21,199	43	76	119
YORO	101,965	41,134	60,831	145	214	359
TOTAL	**1,433,664**	**697,949**	**735,715**	**2,516**	**2,530**	**5,046**

VII. Flexibilidad y Ventajas del Marco

Indudablemente la elaboración de un "Marco de Muestreo" que satisfaga las necesidades de diferentes encuestas es una tarea permanente y de estrecha coordinación con las unidades de Censos y Encuestas, Cartografía y la oficinas de Computo que son las encargadas de almacenar la información (proveniente de los censos y encuestas que se realizan en el país).

El presente Marco tiene la flexibilidad necesaria como para poder responder a otro tipo de necesidades de información que no requieran un nivel de desagregación, que es el sector censal los cuales están organizados jerárquicamente y geográficamente de manera que el marco puede utilizarse con cualquiera de los propósitos siguientes:

- Seleccionar directamente las UPMs.
- Poder formar nuevas UPMs, de cualquier tamaño deseado agrupando los sectores adyacentes.
- Su construcción esta realizada de manera tal que permite una diversidad de **unidades de muestreo,** para diferentes encuestas que tengan necesidad de información diferente de los hogares. Dado que las mismas pueden requerir UPMs más pequeñas que los sectores censales o quizás de mayor tamaño en especial cuando se trata de investigar una característica poco frecuente en la población.
- Se puede seleccionar muestras de subdivisiones administrativas que así se requieran, extrayendo muestras en una o más etapas adicionales si el diseño muestral así lo requiere.
- Evita la duplicidad de esfuerzos y gastos en relación a los trabajos de actualización y recopilación de materiales de muestreo necesario para la selección de la muestra, lo que implica que pueden emplearse más recursos y esfuerzos para mejorar el diseño de la muestra mejorando la calidad de los mapas. La segmentación y otros materiales necesarios para la encuesta, ya que los costos pueden distribuirse entre las diferentes encuestas.
- Permite una selección rápida y económica de diferentes muestras para distintas encuestas controlándose mejor las operaciones de campo entre las encuestas o diferentes rondas de la misma encuesta.
- Facilita el empleo de los mismos materiales de muestreo por diferentes encuestas abaratando de este modo los costos de ellas.

VIII. Actualización del Marco Muestral

Una de las actividades más importantes después de la elaboración del Marco Muestral, es establecer un plan de mantenimiento y actualización cartográfica. Si bien esta es una tarea bastante costosa y que por lo general una sola encuesta no puede ejecutar una actualización completa del marco, esta labor puede ser compartida con otras encuestas que emplearan el marco.

La actualización del Marco Muestral, permite la ubicación exacta de todas las unidades de investigación seleccionadas permitiéndonos lograr una cobertura completa de la encuesta.

La actualización consiste en detectar los cambios debidos al aumento o disminución del número de viviendas, en la UPM como también detectar los cambios de límites geográficos debido a estos cambios.

Este plan de actualización debe desarrollarse en dos fases distintas y complementarias entre si y que son las siguientes:

- **Primera Fase**

  Esta primera fase debe ser la de completitud del marco, se refiero esto a que una vez obtenida toda la información del Censo Nacional de Población y Vivienda el presente marco debe completarse y actualizarse en función de esta información, introduciendo la misma a la base de

datos del marco muestral, de tal manera que la estructura del marco ya no sólo contendrá información del número de viviendas sino que se tendrá toda la estructura dada en los incisos (c) y (d) de la pagina 11.

- **Segunda Fase**

  Esta es una actividad especifica de trabajo de campo en la que ya se tiene que actualizar cartográficamente en el terreno, las unidades de primera y segunda etapa, para evaluar los cambios que se presentaron dentro de las UPM y SCs, que permita realizar las correcciones necesarias para la obtención de resultados siempre óptimos para la encuesta.

INFORME TÉCNICO DE LAS EXPERIENCIAS DE MÉXICO

INSTITUTO NACIONAL DE ESTADÍSTICA GEOGRÁFICA E INFORMÁTICA
(INEGI)

ÍNDICE

Página

A. Los Censos de Población y Vivienda y los Marcos de Muestreo Utilizados en las Encuestas en Hogares 221
B. Diseño y Construcción del Marco de Muestreo 222
B.1 Dominios de estudio 222
B.2 Formación de unidades de muestreo 222
C. Aspectos Estadísticos Asociados a la Construcción del Marco de Muestreo 223
C.1 Estratificación 223
C.2 Tamaño de muestra 223
C.3 Afijación de la muestra 224
C.4 Selección de la muestra 224
C.5 Estimadores 226
C.6 Estimación de las precisiones 227
D. Aspectos Operativos para la Construcción del Marco de Muestreo 229
1. Procedimientos operativos para la elaboración estratificación 229
2. Procedimientos para la actualización de las áreas de listado 233
3. Actualización permanente del marco de muestreo y de la muestra 234
4. Registro de los cambios de las viviendas en la columna 6 ("razón y fecha en la que se suprimió"), del registro de listado de viviendas 236
Construcción y Sistematización del Marco Nacional de Viviendas 240
I. Situación del Marco Muestral 240
II. Insumos 240
III. Cobertura Geográfica 240
IV. Formación de Unidades de Muestreo (UPM) 240
Algoritmo para la formación de UPM en áreas urbanas 241
Estratificación de las UPM 242
Diseño en dos etapas 243
Afijación de la muestra y selección de UPM 243
Elaboración de listados de viviendas 243
Sistematización de los listados de viviendas 243
Anexos 246

A. Los Censos de Población y Vivienda y los Marcos de Muestreo Utilizados en las Encuestas en Hogares.

En México el diseño y planeación de los censos de población y vivienda y los marcos de muestreo utilizados en las encuestas en hogares, tienen como insumo fundamental el marco geoestadístico nacional (MGN), en un proceso dinámico en el que los levantamientos censales y de encuestas se alimentan de la información registrada en el MGN y a su vez, proporcionan la información para su actualización.

Marco Geoestadístico Nacional Es el sistema constituido por el conjunto de definiciones y reglas de asociación en materia de delimitación de áreas geográficas, así como de los materiales de apoyo para su identificación, que permiten relacionar la información estadística con el área geográfica correspondiente. Se tienen definidas las siguientes áreas geográficas:

Área Geoestadística Estatal (A.G.E.E.) Es el Area Geográfica que contiene a todos los municipios pertenecientes a una Entidad Federativa, por lo cual el país está conformado por 32 AGEE. A cada una de ellas se les ha asignado una clave consecutiva, formada por dos números, de acuerdo al orden alfabético de los nombres oficiales de las Entidades Federativas.

Área Geoestadística Municipal (A.G.E.M.) Es el Area Geográfica que contiene a todas las localidades pertenecientes a un municipio, son el producto de la división de la AGEE, a las cuales se les ha asignado una clave compuesta por tres dígitos, no siempre consecutivo, ni necesariamente de acuerdo al orden alfabético de sus nombres.

Área Geoestadística Básica (A.G.E.B.) Estas Areas Geoestadísticas constituyen la Unidad Básica del Marco Geoestadístico; sus límites están representados por rasgos naturales o culturales, todos ellos reconocibles y perdurables en el terreno.

Se dividen en Urbanas y Rurales, asignándose a cada una de ellas una clave compuesta por tres números, un guión y un número o letra llamado dígito verificador.

AGEB Urbana. Es un conjunto de manzanas (generalmente menor a 50), perfectamente delimitado por calles, avenidas y/o rasgos naturales.

Estas se encuentran en todas las localidades de 2,500 habitantes o más y en las cabeceras municipales que no cumplen con el mínimo de habitantes para ser consideradas urbanas.

AGEB Rural. Porción Geográfica de aproximadamente 10,000 hectáreas que puede incluir localidades no Urbanas que van desde caseríos dispersos hasta localidades que cuentan con menos de 2,500 habitantes.

En el interior de un AGEB Rural puede haber una o más AGEB Urbanas perfectamente bien delimitadas.

El marco muestral de propósitos múltiples del INEGI.

El Marco Muestral de Propósitos Múltiples del INEGI es una muestra maestra obtenida del marco geoestadístico nacional que fue actualizado en su información demográfica y cartográfica a partir del conteo de población y vivienda de 1995.

Los materiales cartográficos utilizados fueron:

- Plano de localidad y croquis individual de AGEB urbanos, los cuales sirvieron para la delimitación de las UPM.
- Carta Topográfica escala 1:50000, sirve para delimitar las AGEB rurales.
- Croquis municipal, en el que son delimitadas todas las AGEB rurales del municipio.

Se cuenta con información estadística y georeferenciada de cada una de las AGEB y tienen las siguientes características:

- Son perfectamente reconocibles en el terreno al estar delimitadas por rasgos topográficos identificables y perdurables.
- Por lo general son homogéneos en cuanto a sus características socioeconómicas.
- Su extensión es tal que pueden ser recorridas por una persona en un día normal de labores.

B. Diseño y Construcción del Marco de Muestreo.

La muestra maestra fue diseñada primordialmente para servir como marco de muestreo de la encuesta nacional de empleo (ENE), que es la encuesta más completa en términos de cobertura geográfica y por tanto con el mayor tamaño de muestra, en relación al resto de encuestas que levanta el INEGI; de aquí que la muestra maestra tiene la suficiencia en cobertura y tamaño de las unidades de muestreo, para constituirse en el marco de muestreo de todas las encuestas en hogares del instituto.

El diseño del marco responde al diseño de muestreo probabilístico con cobertura nacional, estratificado, polietápico y por conglomerados, donde las diferentes unidades de muestreo y los dominios de estudio se describen a continuación:

B.1. Dominios de estudio

Al interior de cada entidad federativa se conforman zonas que agrupan a las localidades de la siguiente forma:

ZONA	DESCRIPCION
Urbano Alto	Áreas metropolitanas, capitales de estado y resto de las ciudades de 100,000 y más habitantes.
Complemento Urbano de alta densidad	Localidades de 20,000 a 99,999 habitantes.
Complemento Urbano de baja densidad	Localidades de 15,000 a 19,999 habitantes.
Rural	Localidades de 2,500 a 14,999 habitantes.
	Localidades con menos de 2,500 habitantes.

B.2. Formación de unidades de muestreo.

Al interior de cada zona la formación de las unidades de muestreo se realiza como a continuación se describe:

Unidad Primaria de Muestreo (UPM):

Las Unidades Primarias de Muestreo están constituidas por una o parte de un Área Geoestadística Básica (AGEB), o de varias AGEB colindantes cuando no se completa el mínimo de viviendas requeridas para formar la UPM. Dependiendo de la zona de referencia su constitución es la siguiente:

UPM en urbano alto.

- Un AGEB con un mínimo de 480 viviendas.
- La unión de dos o más AGEB contiguas y del mismo estrato, con un mínimo de 480 viviendas en conjunto.

UPM en el resto de las zonas.

Un AGEB o la unión de dos o más AGEB que contengan:

- De 100 a 1,000 viviendas en Localidades Urbanas.
- De 50 a 850 viviendas en Localidades Rurales.

Unidad Secundaria de Muestreo (USM)

USM en urbano alto.

La formación de la Unidad Secundaria de Muestreo (USM) o Área de Listado, se realiza sólo en las ciudades ENEU[1] agrupando las viviendas bajo las siguientes condiciones:

- Una manzana que tenga un mínimo de 40 viviendas habitadas.
- Dos o más manzanas contiguas con al menos 40 viviendas habitadas.

USM en el resto de las zonas.

La Unidad Secundaria de Muestreo (USM) está constituida por las viviendas particulares, habitadas permanentemente o aptas para habitarse.

Unidades Terciarias de Muestreo (UTM):

Las Unidades Terciarias de Muestreo (UTM) se tienen solamente en la zona urbano alto ENEU y están constituidas por las viviendas particulares, habitadas permanentemente o aptas para habitarse en el momento de la elaboración de los listados.

C. Aspectos Estadísticos Asociados a la Construcción del Marco de Muestreo.

C. 1 Estratificación.

Al interior de cada entidad y zona, que constituye una primera estratificación geográfica, las UPM se sometieron a una estratificación en términos socioeconómicos utilizando las siguientes variables:

- Porcentaje de población de 6 a 14 años que sabe leer y escribir.
- Porcentaje de población de 15 años y más alfabeta.
- Porcentaje de viviendas con drenaje conectado a la red pública.
- Porcentaje de viviendas con agua entubada dentro de la vivienda.

C.2 Tamaño de muestra

El tamaño de cada muestra expresado en términos de proporción se calcula para una variable básica considerada de interés primordial para la encuesta correspondiente, fijando una confianza y un máximo error esperado, y considerando la experiencia obtenida respecto al efecto de diseño del marco para esa variable y la tasa de no respuesta esperada; mediante la siguiente expresión:

[1] 48 Ciudades autorepresentadas que constituyen el domino de estudio de la Encuesta Nacional de Empleo Urbano.

$$n = \frac{z^2 \text{ q DEFF}}{r^2 \text{ p } (1\text{-TNR})}$$

Donde:

N = Tamaño de la muestra.
P = Estimación de la proporción.
q = (1-p).
r = Error relativo máximo aceptable.
z = Valor asentado en las tablas estadísticas que garantiza realizar las estimaciones con una confianza prefijada.
DEFF = Cociente de la varianza del diseño utilizado entre la varianza del muestreo aleatorio simple para un mismo tamaño de muestra.
TNR = Tasa de no respuesta esperada.

C.3 Afijación de la muestra.

La afijación de la muestra se realiza dentro de cada entidad entre los diferentes estratos de manera proporcional a su tamaño, para lo cual se emplea la siguiente expresión:

$$n'_{gh} = \frac{N'_{gh}}{N'_g} n'_g$$

Donde:

n'_{gh} = Número de viviendas en muestra para el h-ésimo estrato en la g-ésima entidad.

n'_g = Número total de viviendas en muestra para la g-ésima entidad.

N'_{gh} = Número total de viviendas en el h-ésimo estrato en la g-ésima entidad.

N'_g = Número total de viviendas en la g-ésima entidad.

C.4 Selección de la muestra.

A efecto de ejemplificar el cálculo del factor de expansión para una encuesta particular, a continuación se presenta el proceso de selección de la muestra de la ENE-2001. La selección de la muestra de la ENE, se realizó de manera independiente en cada entidad - estrato, el procedimiento de selección varía dependiendo de la zona.

Urbano Alto (Selección trietápica)

1. Se seleccionaron n UPM, con probabilidad proporcional al número de viviendas en el estrato.
2. En cada UPM se seleccionaron k áreas de listado de las 6 seleccionadas para el marco de propósitos múltiples (MPM), utilizando muestreo sistemático con arranque aleatorio.
3. En cada área de listado seleccionada, se seleccionaron 6 viviendas con igual probabilidad.

En consecuencia, la probabilidad de seleccionar una vivienda es:

$$P\{v_{ghijk}\} = \frac{n_{gh} \cdot m_{ghi}}{m_{gh}} \cdot \frac{6m_{ghij}}{m_{ghi}} \cdot \frac{k}{6} \cdot \frac{6}{m^*_{ghij}} = \frac{n_{gh} \cdot m_{ghij} \cdot k \cdot 6}{m_{gh}\ m^*_{ghij}}$$

Donde:

n_{gh} = Número de UPM seleccionadas en el h-ésimo estrato en la g-ésima entidad.

m_{ghi} = Número de viviendas de la i-ésima UPM en el h-ésimo estrato de la g-ésima entidad.

m_{gh} = Número de viviendas en el h-ésimo estrato de la g-ésima entidad.

m_{ghij} = Número total de viviendas en la j-ésima USM de la i-ésima UPM en el h-ésimo estrato de la g-ésima entidad.

k = Número total de USM seleccionadas para la ENE-2001.

m^*_{ghij} = Número total de viviendas en la j-ésima USM de la i-ésima UPM en el h-ésimo estrato para la g-ésima entidad al momento del levantamiento ENE-2001

Su factor de expansión está dado por:

$$F_{ghi} = \frac{m_{gh} \cdot m^*_{ghij}}{n_{gh} \cdot m_{ghij} \cdot k \cdot 6}$$

Complemento Urbano (Selección bietápica).

1. Entre las n_{gh} UPM que se seleccionaron en el MPM con probabilidad proporcional a su tamaño, se seleccionaron n^*_{gh} UPM con igual probabilidad.
2. En cada UPM seleccionada, se seleccionaron 50 viviendas con igual probabilidad utilizando muestreo sistemático.

Por lo tanto, la probabilidad de seleccionar una vivienda es:

$$P\{v_{ghij}\} = \frac{n_{gh} \cdot m_{ghi}}{m_{gh}} \cdot \frac{n^*_{gh}}{n_{gh}} \cdot \frac{50}{m^*_{ghi}} = \frac{m_{ghi} \cdot n^*_{gh} \cdot 50}{m_{gh} \cdot m^*_{ghi}}$$

Donde:

n_{gh} = Número de UPM seleccionadas en el h-ésimo estrato en la g-ésima entidad.

m_{ghi} = Número de viviendas de la i-ésima UPM en el h-ésimo estrato de la g-ésima entidad.

m_{gh} = Número de viviendas en el h-ésimo estrato de la g-ésima entidad.

m^*_{ghi} = Número total de viviendas en la i-ésima UPM en el h-ésimo estrato para la g-ésima entidad al momento del levantamiento ENE-2001.

n^*_{gh} = Número de UPM seleccionadas para la ENE-2001 en el h-ésimo estrato para la g-ésima entidad.

Su factor de expansión está dado por:

$$F_{ghi} = \frac{m_{gh} \cdot m^{*}_{ghi}}{m_{ghi} \cdot n^{*}_{gh} \cdot 50}$$

Rural.

1. Entre las n_{gh} UPM que se seleccionaron en el MPM con probabilidad proporcional a su tamaño, se seleccionaron n^{*}_{gh} UPM con igual probabilidad.
2. En cada UPM se seleccionaron 5 segmentos de 10 viviendas aproximadamente, con igual probabilidad.

Por lo tanto, la probabilidad de seleccionar una vivienda está dada por:

$$P\{v_{ghij}\} = \frac{n_{gh} \cdot m_{ghi}}{m_{gh}} \cdot \frac{n^{*}_{gh}}{n_{gh}} \cdot \frac{10 \cdot 5}{m^{*}_{ghi}} = \frac{m_{ghi} \cdot n^{*}_{gh} \cdot 50}{m_{gh} \cdot m^{*}_{ghi}}$$

Donde:

n_{gh} = Número de UPM seleccionadas en el h-ésimo estrato para la g-ésima entidad.

m_{ghi} = Número de viviendas de la i-ésima UPM en el h-ésimo estrato para la g-ésima entidad.

m_{gh} = Número total de viviendas en el h-ésimo estrato para la g-ésima entidad.

n^{*}_{gh} = Número de UPM seleccionadas para la ENE-2001 en el h-ésimo estrato para la g-ésima entidad.

m^{*}_{ghi} = Número total de viviendas en la i-ésima UPM en el h-ésimo estrato para la g-ésima entidad al momento del levantamiento de la ENE-2001.

Su factor de expansión está dado por:

$$F_{ghi} = \frac{m_{gh} \cdot m^{*}_{ghi}}{m_{ghi} \cdot n^{*}_{gh} \cdot 50}$$

C.5 Estimadores

El estimador del total de la característica X:

A nivel estatal es:

$$\hat{X}_g = \sum_h \sum_i \sum_j \sum_k F^u{}_{ghijk} \left(\sum_k x^u{}_{ghijkl}\right) + \sum_h \sum_i \sum_k F^u{}_{ghik} \left(\sum_l x^R{}_{ghikl}\right)$$

Donde:

F^{u}_{ghijk} = Factor de expansión de la k-ésima vivienda en la j-ésima USM de la i-ésima UPM en el h-ésimo estrato para la g-ésima entidad en el área más urbanizada.

X^{u}_{ghijkl} = Valor observado de la característica de interés X en la l-ésima entrevista de la k-ésima vivienda de la j-ésima USM de la i-ésima UPM en el h-ésimo estrato para la g-ésima entidad en el área mas urbanizada.

F^{R}_{ghijk} = Factor de expansión de la k-ésima vivienda en la i-ésima UPM del h-ésimo estrato en la g-ésima entidad del área menos urbanizada.

X^{R}_{ghijkl} = Valor observado de la característica X en la l-ésima entrevista de la k-ésima vivienda en la i-ésima UPM del h-ésimo estrato en la g-ésima entidad para el área menos urbanizada.

A nivel nacional:

$$\hat{X}_{NAL} = \sum_{g=1}^{32} \hat{X}_g$$

Para la estimación de proporciones, tasas y promedios se utiliza al estimador de razón:

$$\hat{R}_{NAL} = \frac{\hat{X}_{NAL}}{\hat{Y}_{NAL}} = \frac{\sum_{g=1}^{32} \hat{X}_g}{\sum_{g=1}^{32} \hat{Y}_g}$$

Donde, la variable $\hat{Y}_{NAL}$ es definida en forma análoga a $\hat{X}_{NAL}$.

C.6 Estimación de las precisiones.

En la evaluación de los errores de muestreo de las principales estimaciones nacionales se emplea el método de conglomerados últimos. El término conglomerado último se utiliza para denotar el total de unidades en muestra de una Unidad Primaria de Muestreo[2] (UPM). En consecuencia, la UPM de la primera etapa de muestreo, es una parte importante en el cálculo de la varianza de estimadores lineales (p. ej. Totales). Así, el Método de Conglomerados Últimos conjuntamente con el de Series de Taylor, son aplicados a los estadísticos no lineales para estimar las precisiones de estos últimos.

La fórmula empleada para calcular la precisión de $\hat{R}_{NAL}$ es la siguiente:

$$\hat{V}(\hat{R}_{NAL}) = \frac{1}{\hat{Y}^{2}_{NAL}} \sum_{g=1}^{32} \left\{ \sum_{h}^{Lg} \frac{n_{gh}}{n_{gh}-1} \sum_{i}^{n_{gh}} \left[\left(\hat{X}_{ghi} - \frac{1}{n_{gh}} \hat{X}_{gh} \right) - \hat{R}_{NAL} \left(\hat{Y}_{ghi} - \frac{1}{n_{gh}} \hat{Y}_{gh} \right) \right]^{2} \right\}$$

Donde:

[2] Véase: Hansen, M.H., W.N. y Madow W.G. , Sample Survey Methods and Theory, (1953), Vol. 1, pag. 242

$\hat{X}_{ghi}$ = Total ponderado de la variable de estudio X para la i-ésima UPM en el h-ésimo estrato de la g-ésima entidad.

$\hat{X}_{gh}$ = Total ponderado de la variable de estudio X para el h-ésimo estrato en la g-ésima entidad.

n_{gh} = Número de UPM en el h-ésimo estrato para la g-ésima entidad.

Estas definiciones son análogas para la variable de estudio Y.
La estimación de la precisión del estimador del total, se calcula con la fórmula siguiente:

$$\hat{V}\left(\hat{X}_{NAL}\right) = \sum_{g=1}^{32} \sum_{h=1}^{L_g} \frac{n_{gh}}{n_{gh}-1} \sum_{i=1}^{n_{gh}} \left(\hat{X}_{ghi} - \frac{1}{n_{gh}} \hat{X}_{gh}\right)^2$$

Las estimaciones de la desviación estándar (D.E.), coeficiente de variación (C.V.) y efecto de diseño se calculan mediante las siguientes expresiones:

$$D.E. = \sqrt{\hat{V}(\hat{\theta})} \qquad C.V. = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}$$

$$DEFF = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\sqrt{\hat{V}(\hat{\theta})_{mas}}}$$

Donde:

$\hat{\theta}$ = Estimador del parámetro poblacional θ

$\hat{V}(\hat{\theta})_{mas}$ = Estimador de la varianza de $\hat{\theta}$ bajo muestreo aleatorio simple.

Finalmente, el Intervalo de Confianza al 90%, está construido de la siguiente forma:

$$I_{1-\alpha} = \left(\hat{\theta} - 1.645\sqrt{\hat{V}(\hat{\theta})},\ \hat{\theta} + 1.645\sqrt{\hat{V}(\hat{\theta})}\right)$$

donde el nivel de significancia es de $\alpha = 0.10$.

En la siguiente tabla se muestran los errores obtenidos para algunas variables en las diferentes encuestas:

Encuesta	Variable	$\hat{\theta}$	C.V.	DEFF
ENE-98	*TDA*	*0.022516*	*0.034514*	*5.791*
ENE-99	*TDA*	*0.017164*	*0.06122*	*2.45620*
ENE-2000	*TDA*	*0.016400*	*0.031695*	*4.01471*
ENE-2001	*TDA*	*0.017099*	*0.035940*	*5.44453*
ENJUV-2000	*Jóvenes con postgrado*	*34,083*	*0.416*	*8.731*
ENIGH-2000	*Prom. de ing. Corr. Total* *Prom. de Gasto Corr. Total*	*23,313.797* *20,685.259*	*0.04378* *0.06057*	*11.9464* *21.3351*

D. Aspectos Operativos para la Construcción del Marco de Muestreo.

A continuación se presenta un resumen del manual de operación:

1. Procedimientos operativos para la elaboración de listados de viviendas.

En oficina se llenarán los rubros de identificación de las formas de registro del listado de viviendas y croquis de manzana y además se localizará el área de listado donde se realizará el trabajo, consultando un plano general de la zona que precise el lugar al que debe dirigirse, así como la ruta más accesible para llegar al área.

Por otra parte, con el fin de facilitar la ubicación en el croquis del AGEB a trabajar, es conveniente conocer los siguientes aspectos que intervienen en la elaboración de un listado.

- Dentro del plano del AGEB se deben identificar la(s) manzana(s) o localidad(es) que integran el área de listado seleccionada.
- Tener presente que en la mayoría de las áreas urbanas, sus límites coinciden con los de una manzana.
- En las zonas rurales, los límites del área de listado coinciden con los de una localidad, sin embargo, algunas veces el área de listado se compone únicamente de una parte de la localidad y en otros casos se conforma de varias localidades.
- Cabe señalar que para facilitar la ubicación e identificación de las áreas de listado a cada una de ellas les fue asignado un número, el cual se debe anotar en el formato de registro de viviendas, en la hoja de croquis y en la cartografía, azurando el área correspondiente.

Una vez realizados los pasos anteriores, se tendrá que dibujar el croquis en la sección "Ruta de Acceso" del listado de viviendas, en donde se señalará el área de listado y la ruta de acceso más fácil.

Finalmente, se deben llenar los datos de identificación de los formatos que se vayan a utilizar. La realización de estas actividades, permitirá contar con los elementos suficientes para salir a campo y localizar el área de listado.

El procedimiento para la elaboración de listados es el mismo, en términos generales, tanto para las áreas de listado urbanas como para las rurales amanzanadas, sin embargo, cada tipo de zona presenta características específicas (conformación y ubicación de las viviendas, rasgos topográficos, etc.) que

requieren de un tratamiento determinado para su registro en los formatos correspondientes, así como de su recorrido sistemático.

Procedimiento en Áreas Urbanas.

Al llegar al área de listado se debe recorrer los límites señalados previamente en el plano para verificar que sean los correctos; si en dicho recorrido de verificación, se da el caso de que los límites del área o al interior de la misma tenga diferencias con el plano, hay que realizar las actualizaciones antes de iniciar el listado del área.

Una vez efectuado el recorrido de los límites y con la seguridad de encontrarse en el área de listado asignada, se inicia la elaboración de listado.

El listado deberá efectuarse de una manera ordenada conforme se van presentando las unidades de vivienda; por lo que se ha establecido un procedimiento para el recorrido del área a trabajar, el cual consiste en iniciar dicho recorrido en la esquina Noroeste ("NO") de cada manzana, después continuar en el sentido en que giran las manecillas del reloj, registrando todas las viviendas que integran la(s) manzana(s) del área de listado. (Ver fig. 39).

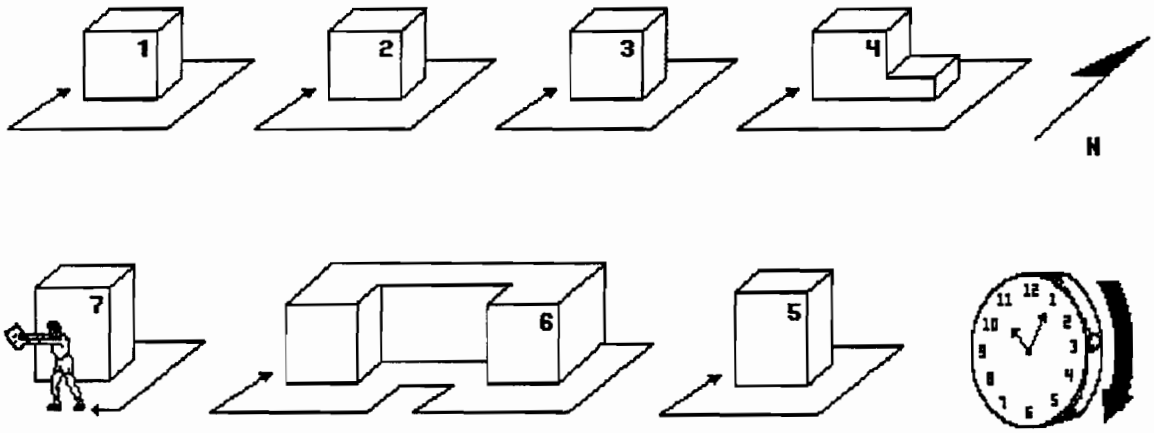


FIG. 39

Por otra parte, cuando la configuración de la manzana sea irregular e impida que éste procedimiento se realice, el listado puede iniciarse en el punto al que primero se llegaría si se hubiera partido de la esquina Noroeste. (Ver fig. 40).

* Punto de inicio

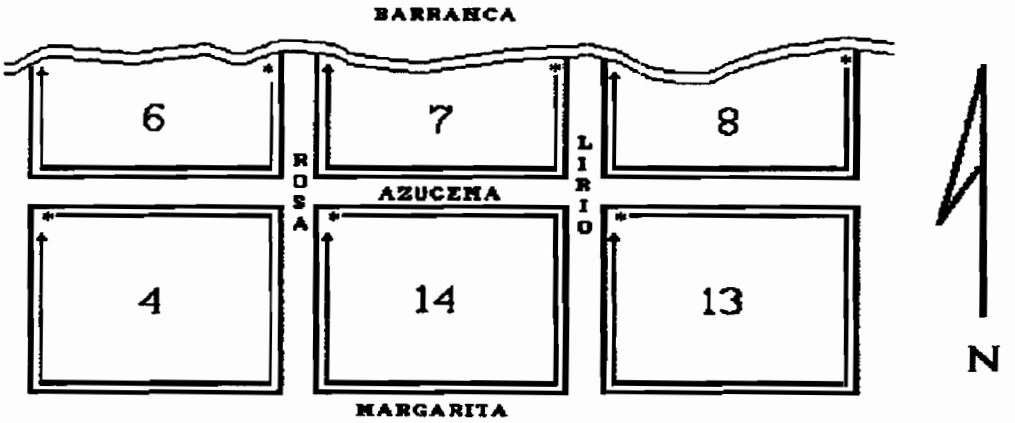


FIG. 40

Hay ocasiones en que la manzana no está bien definida. El criterio a seguir en estos casos, es considerarlas como si fuesen una sola manzana y apegarse lo más posible al procedimiento del recorrido ya descrito. (Ver fig. 43).

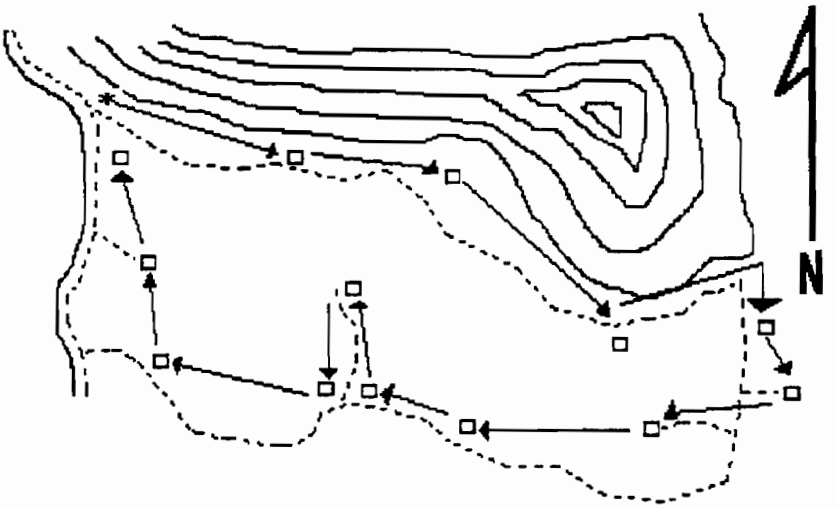


FIG. 43

Es común la presencia de construcciones que tienen más de una vivienda (edificios, vecindades, etc.), éstas se deberán listar conforme vayan apareciendo en el recorrido sistemático.

Existen viviendas que cuentan con más de una entrada, ya sea por la misma o diferente calle, Lo que puede ocasionar que se liste dos veces la misma vivienda; cuando se presente esta situación, hay que registrarla por la entrada principal y en la columna "Descripción o Ubicación de la Vivienda", anotar que cuenta con dos o más entradas.

Al final del documento se anexan el formato de listado de viviendas y del croquis respectivo utilizados en el MNV, debidamente requisitados como ejemplo.

Procedimiento en Áreas Rurales.

Debido a las características específicas de las zonas rurales, la manera de trabajarse en ellas difiere de las urbanas.

En primer lugar, se debe realizar una verificación de los límites del área de listado.

En segundo lugar, se procederá a la elaboración del listado, siguiendo el recorrido sistemático que se haya determinado en función del acceso y de las características del terreno, registrando las viviendas conforme van apareciendo.

En ocasiones las áreas de listado rurales son localidades amanzanadas, para éstas áreas si no se cuenta con un plano de la localidad es necesario elaborarlo y asignarle numeración a las manzanas.

Cuando no exista un trazo uniforme en las calles y caminos, se deberá realizar un plan de recorrido tal y como se había proyectado, sin embargo, hay que tomar en cuenta las siguientes observaciones, ello propiciará realizar una cobertura completa de la zona:

- Partiendo de un punto conveniente, se deberá cubrir cada porción del terreno que se encuentre dentro del área de listado.
- En términos generales, el listado se va realizando a lo largo del camino principal hasta encontrar la intersección con otros.
- Si se encuentran caminos secundarios, se tienen que listar todas las unidades que se localicen en ellos y después, continuar el recorrido por el camino principal.
- Para listar las viviendas, hay que cubrir ambos lados del camino procurando seguir por la vía más accesible, evitando duplicidad en su registro. Además, se tiene que investigar si existen otras unidades de vivienda dispersas dentro del área, aún cuando no se puedan apreciar brechas o veredas que lleguen a ellas.

En virtud de que los predios de zonas rurales pueden estar conformados por una o varias construcciones y que éstas por lo general no están identificadas por números o letras, es preciso que se apegue a las reglas para registrarlas en base al orden establecido, cuando se presente esta situación, unir todas las construcciones que se encuentren dispersas en el croquis y contarla como una vivienda particular, asignando solo un número a uno de los cuartos (ver fig. 45).

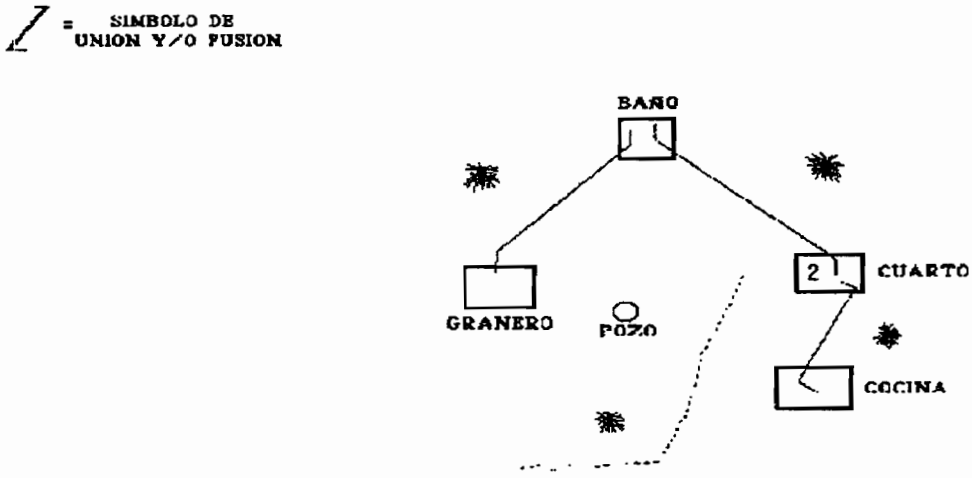


FIG. 45

Revisión y entrega de los listados de viviendas.

Una vez que se haya concluido el listado y antes de abandonar el área revisar lo siguiente:

- Que los rubros de identificación de todos y cada uno de los formatos utilizados cuenten con su respectiva información.
- Tener la seguridad que la extensión total del área del listado haya sido cubierta.
- Que las modificaciones y/o correcciones efectuadas en los croquis se hayan marcado correctamente con color rojo.
- Verificar que el número de vivienda representado en el croquis, sea el mismo que el número de renglón utilizado en la columna "Número progresivo" de la forma "Registro de Listado de Viviendas" MNV 02.
- Por último, asegurarse de haber seguido al pie de la letra las instrucciones del Manual.

Sólo cuando todos los pasos anteriores hayan sido cumplidos, se podrá abandonar el área, regresar a la oficina y entregar al supervisor el material completo.

2. Procedimientos para la actualización de las áreas de listado.

El procedimiento a seguir para efectuar la correcta actualización del listado de viviendas, contiene las siguientes etapas:

- Verificación de los datos de identificación del área de listado.
- Localización del área de listado en oficina y en campo.
- Registro de las modificaciones sufridas por cada vivienda en el listado y actualización simultánea de los croquis anexos.

Estas etapas se efectúan en oficina y en campo, a continuación se describen en detalle, las actividades a realizar:

1. Verificación de los datos de identificación del área de listado.

Una vez suministrado el material, el supervisor recibe su carga de trabajo y como primera actividad, debe verificar los datos de identificación que se encuentran en la parte superior de la carátula de cada listado.

2. Localización del área de listado en oficina.

Cada listado contiene en su carátula una "Ruta de Acceso" al área de listado en cuestión, así como una serie de observaciones, que indican cómo ubicarla para llegar fácilmente y en el menor tiempo posible. Sin embargo, si esta información es dudoza e imprecisa, se debe consultar el plano general de la zona y con base en él, realizar las correcciones y observaciones de la siguiente manera:

- Modificar los datos erróneos que presenta la ruta de acceso con relación al plano general en el listado original o en su defecto, en una hoja blanca si se puede, dibujar la nueva ruta y anexarla al listado.
- Verificar en la carátula del listado, en la sección de "Observaciones del entrevistador", que existan las especificaciones necesarias de cómo llegar al área (ya sea por medio de transporte colectivo o de cualquier otro tipo como: lanchas, caballos, etc.). Por otro lado, si en el momento de llegar a campo, la descripción indicada no coincide con la realidad, actualizar la información en la carátula.

Una tarea que no hay que omitir al hacer los cambios anteriores es lo concerniente a la anotación de los rubros de identificación en la parte superior de la forma de "Registro de Listado de Viviendas".

a) Entidad

b) Municipio

c) Colonia o Localidad

d) No. de Control

e) No. de AGEB

f) No. de Area de Listado

g) No. de Manzana

h) Estrato

i) Distribución

Actividades en campo

1. Localización del área de listado en campo.

Para llegar a la zona, el supervisor tuvo que haber comprobado previamente la efectividad de la ruta de acceso indicada en el listado, si está no cumpliera con su finalidad, se tienen que hacer las correcciones necesarias tal y como se indicó en su oportunidad. Posteriormente, debe identificar y recorrer los límites del área señalados en su croquis, para asegurarse que está ubicado en la zona correcta, a ese respecto, es muy importante señalar que si el supervisor descubre que el listado no se realizó en el área seleccionada, debe reportarlo inmediatamente - vía telefónica - a las oficinas centrales a fin de que se giren las instrucciones pertinentes.

Por otro lado, durante el recorrido del área de listado es posible que el supervisor se encuentre con ciertos cambios o modificaciones ocurridas en el área, con el objeto de contar con los elementos suficientes para solucionar los problemas que surjan a continuación se indican algunos casos detectados con mayor frecuencia, así como sus criterios de solución.

2. Cambio en el nombre de las calles.

Antes de verificar si existen cambios en los nombres de las calles, es importante asegurarse de que se encuentra dentro de los límites del área de listado tal y como aparecen originalmente en el croquis, sin importar que los cambios producidos con posterioridad sean de naturaleza estructural, de nombre o ubicación de una(s) calle(s).

Si el(los) nombre(s) de la(s) calle(s) ha(n) cambiado, el procedimiento a seguir, en todos los casos, consiste en cancelar, con una línea ondulada, el(los) dato(s) original(es) y, debajo de él (ellos), escribir el(los) actual(es) con lápiz rojo.

Cabe señalar la importancia que tiene, para fines de comparación con los archivos de Marcos Estadísticos de oficinas centrales, el no borrar la información original que ha sido objeto de cambios o modificaciones.

3. Fusión de áreas de listado.

Cuando por cambios estructurales, la manzana de una área de listado se haya fusionado con otra, se debe diseñar otro croquis especificando sus nuevos límites y levantar el nuevo listado correspondiente a la manzana fusionada.

4. Desaparición del área de listado.

Si durante su recorrido, el supervisor encuentra que el área de listado ha desaparecido totalmente, ya sea por demolición o por cualquier otra causa, debe hacer las observaciones correspondientes en el listado y deberá reportar esta situación a las oficinas centrales a fin de que se proporcione la sustitución del área y en consecuencia se levante un nuevo listado.

3. Actualización permanente del marco de muestreo y de la muestra.

La actualización del marco de viviendas se lleva a cabo debido a las continuas variaciones de la población y comprende los siguientes tipos de actualización:

Actualización debido a la rotación de la muestra ENE en sus viviendas

Se lleva a cabo cada trimestre, y se actualiza sólo el 20% de la muestra, teniendo al final de cinco trimestres el marco de muestreo completamente actualizado.

A continuación se presentan los principales aspectos de actualización.

1. Áreas de nueva creación dentro de una AGEB urbana.

En este caso se debe a manzanas o construcciones de nueva creación en terrenos que anteriormente eran baldíos, o bien ya existían pero no estaban representados en los planos como consecuencia del crecimiento urbano hacia el interior del AGEB.

En este punto en particular se actualiza su cartografía y se realiza el conteo de viviendas de toda el área, y si el peso de viviendas supera las 900, se realiza una subsectorización; de los subsectores resultantes se selecciona uno, para realizar sus listados correspondientes. Esto se hace con el fin de facilitar el trabajo de campo.

2. Áreas de nuevo crecimiento.

Se considera como área de nuevo crecimiento aquellas que debido al desarrollo urbano de las ciudades, un AGEB rural circundante a la mancha urbana adquiere las características propias de un AGEB Urbano.

Para realizar la actualización tanto cartográfica como en sus viviendas es importante llevar a cabo los siguientes criterios:

- Realizar un recorrido sistemático alrededor de la mancha urbana en su totalidad. El recorrido se iniciará desde un punto bien definido ubicado al norte de la mancha urbana, continuando en el sentido de las manecillas del reloj.
- Se registrarán sobre el mapa con lápiz, las posibles áreas de nuevo crecimiento urbano, al mismo tiempo que se registre un área se numerará en orden progresivo (de la 1 a la "n"), por cada AGEB rural, a fin de llevar un control preciso sobre ellas durante la verificación de campo.
- Verificar en campo las zonas de crecimiento urbano y obtener sus características socioeconómicas (estrato) al que pertenece dicho crecimiento.
- Establecer un programa de recorrido, antes de salir a la verificación de campo buscando ahorrar tiempo, recursos y garantizar la cobertura total. Durante el recorrido verificar y ubicar geográficamente las áreas indicadas y numeradas anteriormente en gabinete y obtener algunas características como nivel de urbanización, conurbanización y nivel socioeconómico, entre otras, para poder clasificarlas, en alguno de los estratos del área metropolitana.
- Trazar los límites del área urbana hasta 1995 y los límites Geoestadísticos, en la cartografía actualizada (Censos económicos 99).
- En dicha cartografía, se vaciará toda la información obtenida en campo, la cual será enviada de inmediato a Oficinas Centrales, donde una vez que se haya utilizado y elaborado una copia de ella, la reintegrará a la Oficina Regional
- En Gabinete ubicar sobre dicha cartografía las zonas nuevo de crecimiento urbano de 1995 a la fecha en los AGEBs rurales y las áreas con mayor factibilidad de crecimiento.

Para formar los sectores que contengan las nuevas áreas de crecimiento urbano se requiere:

a) Delimitar los sectores de reciente crecimiento urbano.

Esto consiste en partir sobre la cartografía con detalles geográficos localizables en el terreno, las nuevas

áreas de crecimiento urbano, de tal modo que cada una de esas particiones sean equivalentes a una UPM y/o un sector, es decir, que tengan 480 viviendas como mínimo.

b) Numeración de sectores.

Estos sectores se localizan dentro de las AGEBs rurales que rodean la ciudad, deberán numerarse y escribir la palabra "sector". considerando que la numeración deberá ser progresiva para cada AGEB rural.

c) Numeración de manzanas dentro de los sectores

Numerar las manzanas, iniciando con el número 901 en adelante para cada uno de los sectores detectados.

d) Delimitar caserío disperso

En los crecimientos como caserío disperso de la AGEB rural, deberán fijarse límites topográficos reconocibles y perdurables en campo, auxiliándose con las cartas topográficas escala 1:50,000.

Después de tener el croquis de dicho sector, se procederá a realizar el conteo exacto de cada manzana que lo forme.

- Informar a las Oficinas Centrales, sobre los avances de las actividades y consulta de dudas y/o problemas en el desarrollo de los mismos.

Con la finalidad de actualizar la muestra de UPMs se lleva a cabo lo siguiente:

a) Incorporación de los sectores a la muestra correspondiente de las áreas de nuevo crecimiento.

La incorporación de las áreas de nuevo crecimiento, se realiza anotando el peso en viviendas de cada sector de la AGEB rural, esto, al final de la muestra en su respectivo estrato, continuando con el acumulado de las viviendas.

b) Selección de UPMs (Sectores)

Se continúa acumulando el intervalo de selección del estrato en cuestión para ver si con la nueva incorporación de viviendas da origen a nuevas selecciones.

4. Registro de los cambios de las viviendas en la columna 6 ("razón y fecha en la que se añadió y suprimió"), del registro de listado de viviendas.

1. Unidades de vivienda agregadas o adicionales con posterioridad a la fecha de elaboración del listado.

Entre los diferentes tipos de viviendas que pueden ser objeto de adición a la forma de Registro de Listado de Viviendas, se encuentran:

a) Viviendas Omitidas y Adicionales. Dentro de este tipo de viviendas se incluyen a las no listadas y de nueva construcción. Con respecto a ellas, puede suceder que formen grandes conglomerados como fraccionamientos, unidades habitacionales, etc. En estos casos, el Supervisor tiene que reportarlo a su jefe inmediato y proporcionarle los siguientes datos:

I. Total de unidades de vivienda que integran el conjunto habitacional.

II. Características de estas unidades (ocupadas, desocupadas, en construcción, uso temporal, etc.).

III. Tipo de construcción (casas o departamentos patrocinados por FOVISSSTE, INFONAVIT, etc., o por particulares).

Si en base a lo anterior se observa que las viviendas constituyen un número considerable, pueden ser objeto de incorporación al Marco Muestral como áreas de nuevo crecimiento, en caso contrario, su registro en el listado debe efectuarse de la siguiente manera:

- Anotar los datos de la nueva vivienda inmediatamente después de la línea horizontal roja que denota la finalización del listado.
- Registrar con lápiz de color rojo en la columna 6 del listado, la clave correspondiente que indica que se trata de una vivienda detectada durante la etapa de actualización.
- Ubicar dicha vivienda en el croquis, tal y como lo indica la Fig. 11 y 11a.

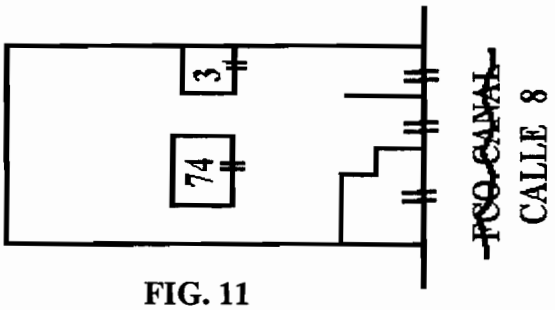


FIG. 11

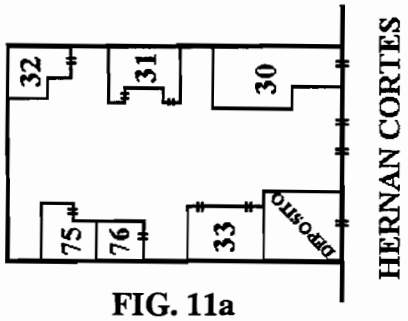


FIG. 11a

También es probable que una vivienda registrada en un principio como unifamiliar, en la etapa de actualización, el Supervisor se encuentre con una vecindad, en este caso, deberá incluir las nuevas viviendas al final del listado y anotar en la columna 6, con lápiz rojo, la clave correspondiente, su nombre y la fecha de actualización.

Si la vivienda en cuestión está seleccionada, se considerará como tal, a la primera estructura que esté ubicada a la derecha y está ocupará el renglón del listado.

Existen también viviendas que a la fecha de la actualización, han cambiado sus características específicas; entre estas se tienen a las desocupadas, viviendas móviles, ocupadas, etc. El proceso de registro en el listado, es el siguiente:

b) Viviendas desocupadas. Si existen viviendas que aparecen registradas como desocupadas, el Supervisor debe verificar, ante todo, la permanencia de su desocupación.

Si efectivamente lo están, no se hace anotación alguna, pero si en el momento de la actualización, la vivienda ya se encuentra ocupada, se deben seguir los siguientes pasos:

- Anular con una línea ondulada de color rojo, la descripción de la vivienda en su carácter de desocupada.
- Después, debajo de la cancelación, anotar con lápiz de color rojo, la nueva condición de la vivienda: "ocupada" - para este caso.

El mismo mecanismo debe llevarse a cabo cuando una vivienda clasificada originalmente como ocupada, en la etapa de actualización, ya está desocupada.

c) Viviendas móviles. Dadas sus particulares características, estas viviendas son susceptibles de cambiar su ubicación en cualquier momento, de aquí que, cuando el Supervisor esté realizando la actualización del listado y se encuentre que:

1. La vivienda se ha trasladado al área seleccionada.

2. La unidad móvil listada previamente, cambió su sitio a otro predio, pero dentro de la misma área, debe agregar ambos casos, al Registro de Listado de Viviendas; asentando su especificación respectiva, el nombre del Supervisor y la fecha de actualización, así como su ubicación en el croquis.

NOTA: En lo que respecta a las unidades de vivienda agregadas con posterioridad a la fecha de elaboración del listado, y a la actualización del croquis, es pertinente señalar tres consideraciones importantes.

- En ningún caso está permitido borrar la información original del listado, ya que esta información hace posible verificar los cambios que se presentan en los listados, al comparar los resultados de la actualización tanto en el listado original como en los archivos de la Oficinas Centrales. En todo caso, lo único permitido es anular (con una línea ondulada) los datos que han cambiado y escribir debajo de ellos la información actualizada. Esta instrucción es válida para todos los rubros del listado, a saber: nombres de calles, características de la vivienda, croquis de vivienda, etc.
- Es importante que cuando el Supervisor agregue viviendas no listadas, respete la instrucción del recorrido del Area de Listado en el sentido de delimitar con los colores correspondientes dicha área, es decir, trazar una línea azul en el listado al finalizar una calle, trazar una línea verde al finalizar una manzana y, por último, trazar una línea roja al finalizar el listado y/o la verificación.
- En lo que respecta al croquis, es importante mantenerlo lo más legible posible, a fin de que la ubicación de las viviendas sea fácil y rápida; por ello, los cambios que sean necesarios anotar, se deben realizar procurando mantener (en todo momento) legible el croquis. Si es necesario, realice croquis anexos, detallando el lugar que le corresponde en el croquis original, así como todos los datos de identificación.

2. Unidades de vivienda a suprimirse.

A diferencia del apartado A, (Unidades de Vivienda Agregadas con Posterioridad a la Fecha de Elaboración del Listado), existen viviendas que están registradas en el listado pero que - por distintos motivos -, deben de ser suprimidas del formato para el Registro de Listado de Viviendas. Estos casos son:

a) Viviendas Demolidas.

Si durante el recorrido de actualización, el supervisor descubre que una vivienda fue demolida o simplemente se trataba de una construcción improvisada con materiales de deshecho y destruida por el deterioro, tiene que seguir los pasos que a continuación se indican:

Escribir con lápiz rojo en la columna 6, la clave "V. D." (vivienda demolida), así como el nombre del Supervisor y la fecha de la actualización.

Por otro lado, puede suceder que el Supervisor descubra que viviendas registradas originalmente en el croquis como demolidas, en la actualidad hayan sido construidas nuevamente. Por lo tanto, el procedimiento a seguir consiste en anotar los datos de estas viviendas después de la línea roja que señala el final del listado.

b) Viviendas convertidas a uso no residencial.

Al actualizar el listado llega a darse el caso de que una estructura utilizada en un principio como unidad de vivienda particular, posteriormente se destine a usos no residenciales como pueden ser bodegas, oficinas, comercios, etc. También es factible que (por equivocación) se haya listado desde un principio este tipo de locales como viviendas.

En ambos casos, el Supervisor tiene que anotar con lápiz de color rojo, en la columna 6 "Razón y Fecha en que se añadió o suprimió", el uso actual de la unidad registrada como vivienda, es

decir, "S" para servicio, "C" para comercio, "B" para bodega, etc. Es necesario anotar, además, la fecha y el nombre del Supervisor.

Por lo que al croquis se refiere, se debe indicar en el espacio destinado a la vivienda, la misma clave que se utiliza en el listado.

c) Vivienda fusionada.

Se dice que hay una fusión de viviendas, cuando se unen dos o más unidades de vivienda para formar una sola, por ejemplo: dos departamentos pueden haberse fusionado para servir de alojamiento a una sola familia.

Ahora bien, la fusión puede darse cuando las personas que ocupan originalmente una vivienda, deciden:

- Ocupar la vivienda contigua, al fondo o ubicada en cualquier otro lugar del mismo predio.
- Ocupar una o más viviendas físicamente separadas, fuera del predio.

Cuando el Supervisor se encuentre con alguno de estos casos, debe tener cuidado en lo siguiente:

Identificar hacia dónde fue el movimiento de fusión. Esto implica conocer cuál vivienda fue ampliada y cuál fue anexada a la primera.

- Si en la fusión, el Supervisor descubre que una vivienda seleccionada, se anexa a una que no lo está, el criterio a seguir será el de considerar a la vivienda que fusiona a las demás, es la que tiene el número progresivo menor; este procedimiento prevalece cuando la fusión se efectúa entre dos o más viviendas seleccionadas.
- Si existen varias unidades de vivienda involucradas en una fusión y una de ellas está seleccionada, y además no se puede determinar el movimiento de fusión; el criterio a seguir será el de considerar a la vivienda seleccionada como aquella que fusiona a las restantes.

Finalmente, en cuanto al procedimiento para el registro de los datos en el listado, se especificará con la ilustración de un ejemplo: supóngase que originalmente una familia ocupaba la vivienda marcada con el número de renglón 69 en el listado; consideremos, asimismo, que posteriormente la misma familia ha ocupado, - además - la siguiente unidad de vivienda (renglón 70 del listado). En este caso, el Supervisor debe realizar lo siguiente:

- Anotar - con un lápiz de color rojo - en la columna "Razón y Fecha en que se Añadió o Suprimió" y a la altura del renglón 70, la leyenda "V. F." al renglón 69 " (Fusionada al renglón 69); además de asentar el nombre del supervisor y la fecha correspondiente.

Procedimiento para la actualización de listados de viviendas en áreas rurales.

En UPM rurales se actualizarán solamente los segmentos que se hayan seleccionado en los Listados de Vivienda, es decir, los espacios geográficos o áreas en que se encuentran comprendidas las viviendas seleccionadas.

Verificación del trabajo.

Por otra parte, antes de abandonar el área de listado en cuestión, el supervisor tiene que revisar:

- Que ha cubierto la totalidad del área.
- Que ha hecho todas las modificaciones o correcciones al listado y al croquis.

Construcción y Sistematización del Marco Nacional de Viviendas

I. Sustitución del Marco Muestral.

En virtud de la realización del censo nacional de población 2000, se cuenta con la información cartográfica y estadística necesaria para la actualización del Marco Muestral actual, sin embargo, es necesaria su sustitución debido al desgaste que ha sufrido por la continua utilización de viviendas requeridas para la rotación de la muestra ENE (20% de la muestra trimestralmente) y para el resto de encuestas continuas y especiales que se han seleccionado de este marco.

II. Insumos.

El censo de población y vivienda del año 2000, provió la cartografía digitalizada en áreas urbanas para la formación automática de las UPM y la información estadística de cada una de las áreas geoestadísticas básicas, que se empleará para calcular las probabilidades de selección y para llevar a cabo la estratificación de las UPM.

III. Cobertura Geográfica.

La cobertura geográfica y organización es similar a la del marco actual.

IV. Formación de Unidades Primarias de Muestreo (UPM)

El tamaño de las unidades primarias de muestreo depende principalmente de la duración del marco, del costo de actualización del marco y de la localización de las viviendas al interior de las UPM. El marco muestral maestro contempla los siguientes dominio de estudio:

a) Urbano Alto, formado por Areas metropolitanas, localidades mayores a 100,000 habitantes y Capitales de Estado.

Los tamaños de cada una de las unidades de muestreo deben ser tales que no representen una carga excesiva en el trabajo de campo cuando se lleve a cabo la actualización del marco muestral. A continuación se dan los tamaños mínimos y máximos por UPM y los criterios para la formación de las unidades de muestreo.

.Cada UPM debe tener un mínimo de 50 viviendas habitadas, un optimo de 80 y un máximo de 160 y esta formado por:

- Una manzana.
- La unión de dos o más manzanas contiguas.
- La unión de dos o más manzanas contiguas de diferentes AGEB colindantes.
- La unión de dos o más manzanas de dos o más AGEBS colindantes de la misma localidad y mismo municipio.
- La unión de dos o más manzanas de dos o más AGEBS colindantes de diferente localidad y mismo municipio.

b) Complemento Urbano, está formado por las localidades de 2,500 a 99,999 habitantes.

1. Cada UPM deberá tener un mínimo de 160 viviendas habitadas, un máximo de 300 y un optimo de 200.
2. Las UPM pueden formarse por:

- Una manzana.
- La unión de dos o más manzanas contiguas.

- La unión de dos o más manzanas contiguas de diferentes AGEB colindantes.
- La unión de dos o más manzanas contiguas de diferentes AGEB colindantes del mismo municipio.

Algoritmo para la formación de UPM en áreas urbanas

Características técnicas:

Este algoritmo esta siendo codificado en el lenguaje de programación Delphi 5, y ha sido probado con cartografía geoestadística a nivel de manzana al cierre del XII Censo General de Población y Vivienda 2000.

Descripción

El algoritmo toma como base la AGEB ubicada más al norte y dentro de ésta la manzana que este situada más al norte, el algoritmo se divide en dos etapas :

Primera etapa:

El algoritmo debe de cumplir con las siguientes restricciones :

- Que las manzanas sean contiguas.
- Las manzanas deben pertenecer a un mismo AGEB.
- El número de viviendas es de 50 , 80 ó 160 (valores que se pueden modificar)

Considerando las restricciones anteriores, el algoritmo comienza a formar UPM de acuerdo a la contigüedad entre manzanas, determinada por el punto de inserción de la clave geoestadística de la manzana y la suma de las cantidades de viviendas correspondientes a cada una, cuidando no sobrepasar el rango establecido.

Cabe mencionar que después de esta etapa es probable que se formen UPM con pesos menores a 50 y mayores a 160, situación que puede ser por las siguientes causas:

- Que la suma de todas las manzanas de una AGEB sea menor a 50.
- Que una manzana por si sola sobrepase el límite superior.

Segunda etapa:

En ésta etapa el algoritmo intenta agrupar UPM de peso por debajo de 50 con otras tal que la suma de las 2 UPM no sobrepase el límite de 160.

Las restricciones son:

- Contigüedad física de UPM
- La distancia que existe entre UPM
- Que las UPM pertenezcan a un mismo municipio

c) Rural, está formado por las localidades menores a 2,500 habitantes.

1. Cada UPM deberá tener un mínimo de 160 viviendas habitadas, un máximo de 300 y un optimo de 200.
2. Las UPM puede constituirse por:

- Un AGEB.
- La unión de dos o más AGEB colindantes.
- Una localidad.

- La unión de dos o más localidades contiguas y del mismo AGEB.
- La unión de dos o más localidades colindantes y de diferente AGEB.

Estratificación de las UPM.

La estratificación de las UPM, se realizará a través de métodos estadísticos multivariados aplicados a los siguientes indicadores del ámbito urbano.

Infraestructura de la vivienda.

Porcentaje de viviendas con agua entubada dentro de la vivienda.
Porcentaje de viviendas con drenaje.
Porcentaje de viviendas con agua entubada, drenaje y electricidad.

Calidad de la vivienda y hacinamiento.

Porcentaje de viviendas con piso diferente de tierra.
Porcentaje de viviendas con paredes de materiales durables.
Porcentaje de viviendas con techos de materiales durables.
Porcentaje de viviendas sin hacinamiento (el cociente "ocupantes entre cuartos", sin contar cocina de uso exclusivo, es menor a DOS).

Equipamiento de la vivienda.

Porcentaje de viviendas con excusado de uso exclusivo que tienen conexión de agua.
Porcentaje de viviendas donde se cocina con gas.
Porcentaje de viviendas con videocasetera.
Porcentaje de viviendas con refrigerador.
Porcentaje de viviendas con lavadora.
Porcentaje de viviendas con teléfono.
Porcentaje de viviendas con calentador de agua o boiler.
Porcentaje de viviendas con automóvil o camioneta propios.
Porcentaje de viviendas con computadora.
Porcentaje de viviendas con refrigerador, lavadora, teléfono y boiler.

Salud.

Porcentaje de la población con derechohabiencia al sistema de salud.

Educación.

Porcentaje de población de 6 a 17 años de edad que asisten a la escuela.
Porcentaje de población de 15 años y más alfabeta.
Porcentaje de población de 15 años y más con postprimaria, incluye carreras técnicas con requisito de primaria terminada o más.
Porcentaje de escolaridad de la población de 15 años y más (grados / población).

Empleo.

Porcentaje de la PEA ocupada femenina.
Porcentaje de la PEA ocupada asalariada (situación en el trabajo) con derechohabiencia.
Porcentaje de la Población Ocupada que percibe más de 2.5 salarios mínimos.
Porcentaje de la Población Ocupada que percibe más de cinco salarios mínimos.
Porcentaje de la PEI de 65 años y más que es jubilada o pensionada.

Porcentaje de la población en hogares pertenecientes a los deciles 9 y 10 según su ingreso per cápita por hogar.
Mediana del ingreso por trabajo del hogar.

Características de los hogares.

Relación de perceptores (población ocupada + población jubilada o pensionada) por cada 10 ocupantes.

Diseño en dos etapas

Se propone para este nuevo marco, emplear un diseño en dos etapas, en áreas urbanas ya que en el área rural se ha venido empleando un diseño bietápico.

El cambio del diseño de tres etapas, a dos etapas se debe a las siguientes consideraciones:

- El costo de traslado de los encuestadores de una UPM a otra no se incrementa.
- Se tiene una muestra de UPM mas dispersa.
- Se obtendrán estratos mas homogéneos al interior dado que las UPM, se forman con unidades mas pequeñas (manzanas).

Afijación de la muestra y selección de UPM

La afijación de la muestra será, en principio, con probabilidad proporcional al tamaño de los estratos mientras que la selección de UPM se hará de manera automática mediante probabilidad proporcional al tamaño, empleando muestreo sistemático con arranque aleatorio.

Elaboración de listados de viviendas

Para la elaboración en campo, se está probando una forma que permita la captura de la información por medio de scanner y así tener una base de datos que permita la selección de las viviendas en forma automatizada, también permitirá que se puedan seleccionar varias encuestas simultáneamente y llevar una adecuada administración del marco muestral de viviendas. A continuación se presenta el proceso de sistematización de los listados así como las formas:

Sistematización de los listados de viviendas

El proceso para la sistematización del Marco Nacional de Viviendas es:

- Diseño y Elaboración de las formas de captura de los listados de viviendas.
- Hacer una prueba de las formas en campo.
- Actualización de los cambios en la forma final.
- Hacer las pruebas de las formas con el paquete **DATA BLOCKS MAGENTA.**
- Elaboración de los listados en cada una de las entidades empleando las formas actualizadas.
- Las formas ya llenas con información de las viviendas, en cada entidad serán enviadas a oficinas centrales.
- En oficinas se procede a hacer la capturas por medio de scanner.
- Evaluación de los resultados de la captura de la información.
- Obtención del Marco Nacional de Viviendas en Base de Datos.
- Implementación de las bases de datos para la selección de las viviendas en forma automatizada.
- De oficina central enviar por e-mail, la selección de las muestras a cada una de las entidades, para levantar las encuestas correspondientes,
- Los cálculos de los factores de expansión se obtendrán en forma automatizada.

HOJA DE

MARCO NACIONAL DE VIVIENDAS
REGISTRO DE LISTADO DE VIVIENDAS

FORMA MNV 02

AREA METROPOLITANA	NÚM.DE CONTROL	ESTRATO	DISTRIBUCIÓN
00 11 22 33 44 55 66 77 88 99	000000 111111 222222 333333 444444 555555 666666 777777 888888 999999	0000 1111 2222 3333 4444 5555 6666 7777 8888 9999	000 111 222 333 444 555 666 777 888 999

CALLE:

ELABORO POR:

fecha:

HOJA ☐ DE ☐

MZA	NÚM. PROG.	CALLE O LOCALIDAD	EDIFICIO	NÚMERO EXT. O DEPTO	NÚMERO INT.	DESCRIPCIÓN O UBICACIÓN DE LA VIVIENDA	TIPO DE VIVIENDA
						Nombre:	

MZA	NÚM. PROG.	CALLE O LOCALIDAD	EDIFICIO	NÚMERO EXT. O DEPTO	NÚMERO INT.	DESCRIPCIÓN O UBICACIÓN DE LA VIVIENDA	TIPO DE VIVIENDA
						Nombre:	

MZA	NÚM. PROG.	CALLE O LOCALIDAD	EDIFICIO	NÚMERO EXT. O DEPTO	NÚMERO INT.	DESCRIPCIÓN O UBICACIÓN DE LA VIVIENDA	TIPO DE VIVIENDA
						Nombre:	

MZA	NÚM. PROG.	CALLE O LOCALIDAD	EDIFICIO	NÚMERO EXT. O DEPTO	NÚMERO INT.	DESCRIPCIÓN O UBICACIÓN DE LA VIVIENDA	TIPO DE VIVIENDA
						Nombre:	

ANEXOS

REGISTRO DE LISTADO DE VIVIENDAS.

FORMA 1 DE 1

INSTITUTO NACIONAL DE ESTADISTICA GEOGRAFIA E INFORMATICA

MARCO NACIONAL DE VIVIENDAS
REGISTRO DE LISTADO DE VIVIENDAS
FORMA MNV

a) ENTIDAD	b) MUNICIPIO O DELEGACION
AGUASCALIENTES 01	AGUASCALIENTES 0011

c) COLONIA O LOCALIDAD(ES): UNIDAD MODELO

d) NUM. DE CONTROL	e) NUM. DE AGEB	U.P.M.
140469	101-4	3

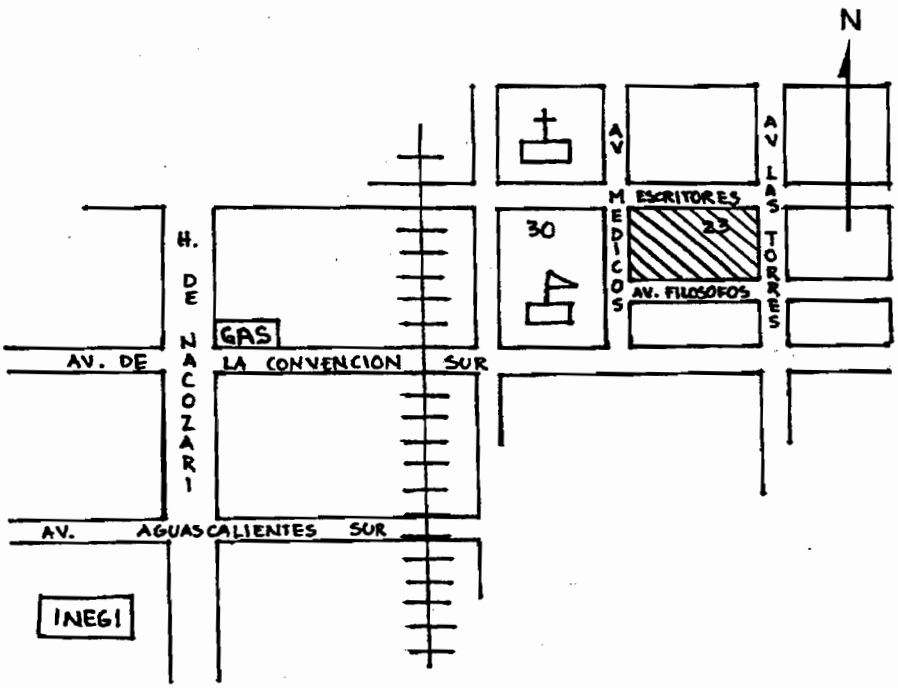
f) NUM. DE AREA DE LISTADO	g) NUMERO DE LA(S) MANZANA(S)	h) ESTRATO	i) DISTRIBUCION
1	23	1403	112

j) FRACCION (F) Y NUMERO ALEATORIO (NA)

F	NA	F	NA	F	NA	F	NA	F	NA

k) ELABORADO POR: RUBEN ATILANO — FECHA: DIA 13, MES 05, AÑO 97

RUTA DE ACCESO AL AREA DE LISTADO



INSTRUCCIONES Y OBSERVACIONES ESPECIALES: TOMAR RUTA 20 EN H. DE NACOZARI Y AV. DE LA CONVENCION HACIA EL ESTE, BAJARSE EN AV. LAS TORRES Y CAMINAR UNA CALLE AL NORTE. AHI SE ENCUENTRA EL AREA DE LISTADO.

ACTUALIZADO POR: FECHA

SR. ENTREVISTADOR: ANOTE CUALQUIER OBSERVACION O ACLARACION QUE SEAN NECESARIAS

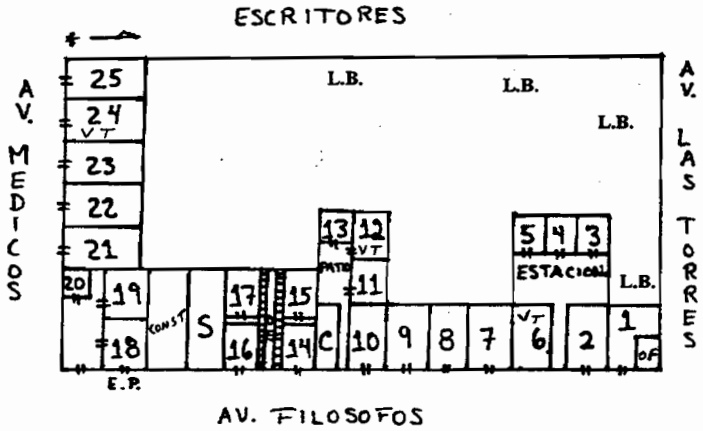
CROQUIS DEL LISTADO DE VIVIENDAS

FORMA 1 DE 1

MARCO NACIONAL DE VIVIENDAS
HOJA DE CROQUIS

a) ENTIDAD: AGUASCALIENTES 01	b) MUNICIPIO O DELEGACION: AGUASCALIENTES 001	
c) COLONIA O LOCALIDAD(ES): UNIDAD MODELO		
d) NUM. DE CONTROL: 140469	e) NUM. DE AGEB: 101-4	U. P. M.
f) NUM. DE AREA DE LISTADO: 1	g) NUMERO DE LA(S) MANZANA(S), CALLE(S): 23	h) ESTRATO: 1403
i) DISTRIBUCION: 112		

NOTA: UBICAR EL NORTE Y EL INICIO DE RECORRIDO



SIMBOLOGIA

C: COMERCIO
CONST. CONSTRUCCION
OF. OFICINAS
S. SERVICIO
V.T. VIVIENDA TEMPORAL
E.P. ENTRADA PRINCIPAL
L.B. LOTE BALDIO
ESTACION: ESTACIONAMIENTO

FORMA MNV-04J

REGISTRO DE LISTADO DE VIVIENDAS.

Número Progr.	Calle o Avenida	Números Ext.	Números Int.	DESCRIPCION O UBICACION DE LA VIVIENDA	Razón y fecha en la que se añadió o suprimió	Para uso exclusivo de la oficina: Tipo de encuesta	Para uso exclusivo de la oficina: Viviendas seleccionadas
1	F	101	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre MARIO GARCIA CUEVAS			
2	I	103	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre JOSE CHAIRES LOPEZ			
3	L	103	1	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre JOEL CRUZ PEREZ			
4	O	103	2	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ANGEL SANCHEZ RUIZ			
5	S	103	3	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ORLANDO CASTRO REYES			
6	O	S/N	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) V.T. Nombre			
7	F	107	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ARTURO GODOY GONZALEZ			
8	I	109	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre JUAN PEREZ GOMEZ			
9	L	111	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre CARLOS VELA MARTINEZ			
10	O	113	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre VERONICA SANTOS ASPE			
11	S	115	1	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre SIN INFORMANTE			
12	F	115	2	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) V.T. Nombre			
13	O	115	3	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre LUIS PEREZ CORREA			
14	S	117	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ANTONIO MACIAS LOPEZ			
15		117	A	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre EFREN LUNA VERA			

REGISTRO DE LISTADO DE VIVIENDAS.

Núm. Preg.	Calle o Avenida	Número Ext.	Número Int.	DESCRIPCION O UBICACION DE LA VIVIENDA	Razón y fecha en la que se añadió o suprimió	Para uso exclusivo de la oficina: Tipo de encuesta	Para uso exclusivo de la oficina: Viviendas entrevistadas
16	FILOSOFOS	121	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ALBERTO GARCIA GUTIERREZ			
17		121	A	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre MARCELA CASTAÑEDA RUIZ			
18		127	1	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre JORGE DIAZ ZEPEDA			
19		127	2	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ANTONIO JIMENEZ HARO			
20		127	3	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ANGEL BUSTOS FERMIN			
21	AV. MEDICOS	202	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre CASTULO OCHOA GUZMAN			
22		204	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre RAUL PADILLA CONTRERAS			
23		206	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre ALFONSO PEREZ VILLALBASO			
24		208	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) V.T. Nombre			
25		210	–	Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre SIN INFORMANTE.			
				Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre			
				Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre			
				Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre			
				Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre			
				Paredes ( ) Techo ( ) Ventana ( ) Puerta ( ) Nombre			

ALGUNOS ASPECTOS RELACIONADOS CON LA MUESTRA DE LAS ENCUESTAS DE HOGARES

ROBERTO CASTILLO

CONTRALORÍA GENERAL DE LA REPÚBLICA
DIRECCIÓN DE ESTADÍSTICAS Y CENSOS
PANAMÁ

ÍNDICE

Página

I.- Situación Sociodemográfica 253
II.- Aspectos Generales de la Encuesta de Hogares 254
III.- Diseño de la Muestra 254
A. Universo o población en estudio 254
B. Marco muestra 256
1. Actualización del marco muestral 256
C. Unidades de muestreo 256
D. Dominio de estudio 256
E. Estratificación 257
F. Tamaño de la muestra 257
G. Selección de la muestra 259
H. Estimación 259
IV.- Resultados de la Encuesta 259

I. Situación Sociodemográfica:

Los resultados finales del X Censo de Población y VI de Vivienda proporcionan información no sólo sobre la composición y estructura de la población, sino que evidencian los avances alcanzados por el país y los retos aún por superar en materia de vivienda, salud, educación, entre otros.

De acuerdo a sus resultados, se puede señalar que nuestro país ha pasado de un crecimiento alto (registrando una tasa de 3.17 entre 1911 y 1920) a un crecimiento más lento (registrando una tasa anual de crecimiento promedio (TCP) de 2.00 entre 1990 y el 2000); situación que según las estimaciones se mantendrá durante los próximos 25 años, como consecuencia directa de la disminución de la fecundidad a nivel nacional.

La población de Panamá pasó durante esta última década de 2,329,329 a 2,839,177 personas, lo que representa un incremento poblacional en diez años de 509,848 personas. El 48.9 por ciento de la población total se concentra en la Provincia de Panamá, generando una TCP de 2.62 por ciento.

En contraposición a lo antes expuesto, se encuentran las provincias de Veraguas, Los Santos y Herrera (con una TCP de 0.52, 0.82 y 0.90 por ciento respectivamente). En el caso de Veraguas, se han logrado progresos en lo concerniente a materia de salud, lo cual ha redundado en una disminución tanto de la fecundidad como de la mortalidad, como consecuencia de la segregación de distritos que en 1990 pertenecían a la misma y que pasaron recientemente a conformar la Comarca Ngöbe Buglé y que se caracterizaban por registrar altas tasas de crecimiento poblacional.

Por otro lado, las provincias de Herrera y Los Santos que tradicionalmente han registrado una baja fecundidad y el desplazamiento de su población hacia otras provincias, parecen mantener esta tendencia, ya que registran TCP de 0.82 y 0.90 respectivamente, incrementando su población absoluta en 8,784 y 6,548.

TASA DE CRECIMIENTO INTERCENSAL POR PROVINCIA: CENSOS 1911 A2000

Provincia	1911-20	1920-30	1930-40	1940-50	1950-60	1960-70	1970-80	1980-90	1990-00
Total	3.17	0.47	2.76	2.56	2.94	3.06	2.51	2.58	2.00
Bocas del Toro	2.03	-5.57	0.39	3.02	3.83	3.12	2.10	3.32	1.87
Coclé	2.86	0.66	1.38	2.69	2.45	2.54	1.75	2.08	1.57
Colón	6.85	-0.19	3.01	1.41	1.58	2.60	2.17	2.32	1.95
Chiriquí	2.11	0.06	3.56	2.15	3.15	2.43	2.00	1.14	1.36
Darién	1.98	2.25	1.04	-0.18	3.01	1.50	1.56	3.06	1.17
Herrera	2.60	0.68	1.97	2.71	2.10	1.74	1.22	1.34	0.90
Los Santos	1.58	1.76	1.78	2.11	1.40	0.27	-0.31	0.91	0.82
Panamá	5.25	1.63	4.05	3.59	4.14	4.76	3.71	2.85	2.62
Veraguas	1.24	0.43	1.92	2.28	2.10	1.52	1.32	1.37	0.52
Comarca Kuna Yala	....	....	....	....	....	....	....	1.75	-0.48
Comarca Emberá	....	.....	.....	....	.....	.....	.....	....	0.34
Comarca Ngöbe Buglé	.....	.....	.....	....	....	.....	....	....	4.27

Cabe destacar, que la disminución se observa al comparar las TCP registradas entre 1980-90 y 1990-2000 de la Provincia de Bocas del Toro (1.87) y la Provincia de Chiriquí (1.36) las cuales fueron las más afectadas al conformar la Comarca Ngöbe Buglé.

En lo que se refiere al comportamiento de la población por sexo, el Censo registró que habían 1,432,566 hombres y 1,406,611 mujeres, generando un índice de masculinidad de 101.8 hombres por cien mujeres lo que implica que habría un hombre adicional por cada cien mujeres, contrario a lo que comúnmente se piensa.

No obstante lo anterior, se encuentran diferencias a nivel de las provincias, registrándose los índices de masculinidad más altos en las provincias de Darién, Veraguas y la Comarca Emberá (con 123.5, 111.2 y

113.6 hombres por cien mujeres) y los más bajos en la Comarca Kuna Yala, Comarca Ngöbe Buglé y en la Provincia de Panamá (87.6, 97.9 y 98.2 hombres por cien mujeres respectivamente).

En cuanto a la estructura por edad, las cifras revelan un envejecimiento de nuestra estructura poblacional ya que mientras en 1990 los menores de 15 años representaban casi el 35% de la población, en el 2000 este grupo de edad representa el 32.2% de la población total; no obstante lo anterior, continúa siendo un porcentaje significativo de población joven que demandará una serie de servicios tanto del sector educativo como del sector salud.

Por otro lado, el 62% de nuestra población tiene de 15 a 64 años lo que implica una gran volumen de población en edad de trabajar y por ende una fuerte presión para el mercado laboral y finalmente el 6% de la población tiene 65 años y más de edad lo que representa al compararla con la cifra obtenida en 1990 un incremento de casi un punto. Estas transformaciones en nuestra estructura han provocado un incremento en la edad mediana de la población pasando este indicador de 22 años en 1990 a 24 años en el 2000.

ESTRUCTURA POR EDAD DE LA POBLACIÓN: CENSOS
1980 A 2000

Grupos de edad	1980	1990	2000
Total	100.0	100.0	100.0
Menores de 15 años	39.1	34.9	32.0
15-64	56.6	59.9	62.0
65 y más	4.3	5.2	6.0

II. Aspectos Generales de la Encuesta de Hogares

La Dirección de Estadística y Censo realiza desde 1963 en forma ininterrumpida las Encuestas de Hogares. La elaboración de la muestra utilizada está basada en los Censos de Población y Vivienda realizados en el 2000 y las estimaciones de Población preliminares con referencia a dichos Censos.

Con el propósito de aumentar la calidad de la información proporcionada, se ha incorporado por primera vez en las Encuestas de Hogares a la población que reside en áreas de difícil acceso y en áreas indígenas. En ese sentido, la encuesta permite un mayor nivel de desagregación geográfica al presentar de forma novedosa información a nivel provincial y de áreas urbanas y rurales, respectivamente. Para la provincia de Panamá se presenta la información desagregada a nivel de Ciudad de Panamá, Resto del Distrito de Panamá, Distrito de San Miguelito, Panamá Oeste y Panamá Este.

La Encuesta de Hogares tiene como Universo a la población de 15 años y más que reside habitualmente en viviendas particulares. Los datos obtenidos en esta encuesta tienen como referencia la semana que antecede a aquella en que se efectúan las entrevistas. En consecuencia, las cifras corresponden a un promedio semanal del mes de agosto.

La ampliación de la cobertura de la Encuesta y las mejoras metodológicas introducidas constituyen evidentemente mayor disponibilidad de información para la caracterización del mercado laboral panameño, entre otros aspectos, que podrá fortalecer los diversos programas e investigaciones en este campo que realizan tanto las instituciones públicas como privadas del País.

III. Diseño de la Muestra:

A. Universo o población de estudio:

El Universo de estudio lo constituye la población de 15 años y más de edad residente en las viviendas particulares ocupadas del país, tanto del área indígena como en la no indígena.

De acuerdo al Censo de Población y Vivienda de mayo de 2000, la población total en el universo ascendió a 2,791,637, la no indígena a 2,601,757 y la indígena a 189,880.

A nivel de la población de 15 años y más, está ascendió a 1,890,162 personas, de las cuales, el 65% corresponde al área urbana y el 35% al área rural. La población no indígena está constituida por 1,794,067 personas y la indígena por 96,095 personas.

VIVIENDAS PARTICULARES Y POBLACION DE 15 AÑOS Y MÁS EN LA REPUBLICA, SEGUN PROVINCIA Y ÁREA : CENSO DE POBLACION 2000

Desagregación geográfica	Vivienda particulares			Poblacion total 1/	Población de 15 años y más
	Total	Ocupadas	Con población de 15 años y más		
Total	790,813	681,803	681,590	2,791,637	1,890,162
Urbano	497,063	442,840	442,743	1,733,225	1,231,220
Rural	293,750	238,963	238,847	1,058,412	658,942
No Indígena	757,596	653,211	653,033	2,601,757	1,794,067
Urbano	497,063	442,840	442,743	1,733,225	1,231,220
Rural	260,533	210,371	210,290	868,532	562,847
Bocas del Toro	15,648	13,718	13,710	66,230	39,136
Urbano	8,573	7,631	7,629	34,173	21,252
Rural	7,075	6,087	6,081	32,057	17,884
Coclé	54,653	44,496	44,480	200,644	130,496
Urbano	16,941	14,229	14,225	60,778	42,577
Rural	37,712	30,267	30,255	139,866	87,919
Colón	57,036	49,715	49,706	200,354	130,964
Urbano	37,897	34,713	34,710	135,151	91,154
Rural	19,139	15,002	14,996	65,203	39,810
Chiriquí	99,379	86,901	86,871	351,569	239,005
Urbano	44,053	39,762	39,752	159,439	112,306
Rural	55,326	47,139	47,119	192,130	126,699
Darién	9,879	7,848	7,832	32,296	18,817
Urbano	397	346	346	1,449	917
Rural	9,482	7,502	7,486	30,847	17,900
Herrera	31,712	27,202	27,200	101,861	72,264
Urbano	14,636	12,887	12,885	47,636	35,277
Rural	17,076	14,315	14,315	54,225	36,987
Los Santos	31,962	25,052	25,049	82,787	61,992
Urbano	6,347	5,310	5,310	17,850	13,726
Rural	25,615	19,742	19,739	64,937	48,266
Panamá	399,945	349,598	349,517	1,361,366	966,760
Urbano	352,669	314,345	314,276	1,218,579	873,198
Rural	47,276	35,253	35,241	142,787	93,562
Veraguas	57,382	48,681	48,668	204,650	134,633
Urbano	15,550	13,617	13,610	58,170	40,813
Rural	41,832	35,064	35,058	146,480	93,820
Indígena	33,217	28,592	28,557	189,880	96,095
Rural	33,217	28,592	28,557	189,880	96,095

1/ Excluye la población que habita en viviendas colectivas y las personas sin vivienda.

En lo relativo a las viviendas particulares el Censo detectó 790,813 viviendas, de las cuales el 63% están ubicadas en el área urbana y el 37% en el área rural. Del total de viviendas particulares, 681,803 viviendas que representan el 86% están ocupadas.

Las viviendas no indígenas ascendieron a 757,596 y las indígenas a 33,217, representando con respecto al total de viviendas particulares el 86% en ambos universos. En cuanto a su composición con respecto al total nacional el universo no indígena representa el 96% y el indígena el 4%.

Cabe destacar que la provincia de Panamá, la cual es el centro donde se concentra la actividad económica, política y social del país; reúne más del 50% de las viviendas y la población del universo de estudio.

B. Marco Muestral:

Con base a la información proveniente del Censo de Población de mayo de 2000 y la segmentación censal proveniente de la actualización cartográfica, se conformó el marco muestral.

Es importante destacar que si bien la base fue la segmentación censal (unidades de empadronamiento), se procedió a evaluar la cantidad de viviendas por segmento tratando de homogenizar la cantidad de viviendas por segmento; estableciéndose para el área urbana un promedio de 12 viviendas y para el área rurald e 10 viviendas.

El total de segmentos censales que constituyó el marco muestral, ascendió a nivel nacional a 69,709; de los cuales 64,901 son no indígenas y 4,808 son indígenas. El universo no indígena contenía un total de 38,857 segmentos urbanos y 26,044 segmentos rurales.

COBERTURA DE SEGMENTOS CENSALES EN LA REPUBLICA, SEGÚN UNIVERSO Y AREA: CENSO DE POBLACION 2000

Desagragación geográfica	Total de segmentos	Distribución relativa
Total	69,709	100
Urbano	38,857	56
Rural	30,852	44
No Indígena	64,901	93
Urbano	38,857	60
Rural	26,044	40
Indígena	4,808	7
Rural	4,808	7

1. Actualización del Marco Muestral:

Una vez realizada la selección de las unidades primarias de muestreo se procede a realizar una gira de actualización cartográfica, en la que se incluye la actualización de las nuevas áreas de desarrollo, asentamientos espontáneos y edificios nuevos de 12 y más viviendas.

Durante la actualización se levanta una planilla donde se registran las viviendas, la condición de ocupación en al que se encuentran y si es posible el nombre del jefe de la vivienda; esta planilla incluye la localización especifica de la vivienda.

C. Unidades de Muestreo:

Las unidades de muestreo constituyen la unidad de selección de la muestra y se representan de la siguiente manera:

- Unidad primaria de muestreo: constituida por el segmento censal.
- Unidad secundaria de muestreo: constituida por las viviendas particulares.

D. Dominio de Estudio:

Considerando el tipo de estudio y la necesidades de los usuarios y tendiendo siempre a una mayor utilidad de la información, los dominios de estudio identificados son los siguientes:

Dominios de estudio de la Encuesta de Hogares

Sector de estudio	Universo	
	No indígena	Indígena
Nacional	X	X
Urbano	X	X
Rural	X	X
Provincia	X	X
Urbano	X	X
Rural	X	x

La provincia de Panamá, donde se ubica la capital del país y se concentra la mayor cantidad de población se subdividió en los siguientes dominios de estudio:

- Distrito de Panamá (Urbano-Rural)
- Ciudad de Panamá
- Resto del Distrito de Panamá (Urbano-Rural)
- Distrito de San Miguelito
- Panamá Oeste (Urbano-Rural)
- Resto de la Provincia de Panamá (Urbano- Rural)

El dominio de estudio indígena a nivel nacional esta integrado Nivel nacional por cada una de las comarcas legalmente constituidas y por las comunidades indígenas fuera de las mismas; que mantiene los mismos patrones socioculturales.

E. Estratificación:

La estratificación asegura un mejor confiabilidad de la muestra al disminuir la varianza de las estimaciones. La teoría del muestreo indica que la estratificación es más eficiente en cuanto más homogéneas sean las unidades que pertenecen a un mismo estrato y más heterogéneas entre sí; en ese sentido la estratificación utilizada en la encuesta cumple con este principio y asegura que cada unidad de muestreo pertenezca a un solo estrato.

La estratificación del universo de estudio esta basada en un criterio geográfico (estratificación implícita), de acuerdo a la codificación político administrativa del país, considerando la división entre áreas. Tal y como se señaló con anterioridad se consideran también en esta fase las áreas en desarrollo, los asentamientos espontáneos y los edificios nuevos de 12 y más viviendas.

La estratificación utilizada ha reflejado hasta estos momentos resultados adecuados en torno a la estructura socioeconómica existente en el país, apoyado principalmente en las particularidades que presenta la distribución geográfica de la población del país.

F. Tamaño de la Muestra:

El cálculo de tamaño de la muestra se realizó utilizando el modelo matemático del muestreo aleatorio simple, calculándose el tamaño para cada provincia a nivel urbano y rural; para la obtención del tamaño muestral se utilizaron como variables criticas: la proporción de la población de 15 años y más de edad y la tasa de desempleo. Se utilizó un nivel de confianza de 95% y un nivel de error que osciló entre 3% y 4% a excepción de la provincia de Darién dada su peso poblacional; para el área urbana de esta provincia se fijó un error muestral de 7% y para el área rural de 5%.

La muestra no contempla reemplazos y se ajustó mediante una tasa de no respuesta de 10% y para las viviendas cerradas de 10%. Se aplicó un efecto de diseño igual a 1.5.

Para la provincia de Panamá, el cálculo de la muestra se obtuvo para cada uno de los dominios de estudio en que se desagregó dicha provincia, contemplando para aquellas que contienen ambas áreas, el cálculo respectivo.

El diseño de la muestra para el área urbana responde a un diseño de dos etapas seleccionándose en la primera la unidad primaria de muestreo y en la segunda la mitad de las viviendas de la unidad primaria de muestreo siendo en promedio 6 viviendas. En la provincia de Darién la muestra urbana se seleccionó en una sola etapa por las razones ya explicadas.

En el área rural, el diseño de la muestra corresponde a un muestreo monoetápico; es decir de una sola etapa, seleccionándose la unidad primaria de muestreo y a lo interno de ella se toman todas las viviendas motivada esta decisión, por los altos costos de la entrevista en esta área. La muestra indígena por ende responde al mismo diseño muestral.

TAMAÑO DE LA MUESTRA, SEGUN PROVINCIA Y AREA

Desagregación geográfica	Unidades de muestreo	Viviendas seleccionadas
Total	2,032	13,108
Urbano	1,292	7,247
Rural	740	5,861
No Indígena	1,976	12,775
Urbano	1,292	7,247
Rural	684	5,528
Bocas del Toro	112	796
Urbano	72	400
Rural	40	396
Coclé	160	1,000
Urbano	92	461
Rural	68	539
Colón	180	1,221
Urbano	100	590
Rural	80	631
Chiriquí	172	1,210
Urbano	100	571
Rural	72	639
Darién	48	413
Urbano	12	122
Rural	36	291
Herrera	156	1,010
Urbano	84	465
Rural	72	545
Los Santos	168	993
Urbano	104	469
Rural	64	524
Panamá	792	4,949
Urbano	628	3,640
Rural	164	1,309
Veraguas	188	1,183
Urbano	100	529
Rural	88	654
Indígena	56	333
Rural	56	333

G. Selección de la Muestra:

El proceso de selección de la muestra para el área urbana con base a la probabilidad que corresponde, se realizó para la primera etapa, mediante la selección sistemática con probabilidad proporcional al tamaño. En la segunda etapa, se seleccionó con probabilidades fijas (medio segmento), seleccionándose sistemáticamente conglomerados de dos viviendas, obteniéndose en promedio tres conglomerados de dos viviendas cada uno.

En el área rural la muestra de unidades primarias de muestreo, se obtuvo sistemáticamente mediante la aplicación del muestreo con probabilidad proporcional al tamaño.

H. Estimación

El proceso de inferencia estadística se realiza por área mediante la aplicación del método de razón, utilizando como variable exógena la población de 15 años y más de edad. La estimación a nivel de la República se obtiene mediante la sumatoria de las estimaciones provinciales; a nivel de provincia mediante las sumatorias a nivel de área urbana y rural

Para la muestra indígena el método se apoya igualmente en el de razón y se aplica a nivel nacional.

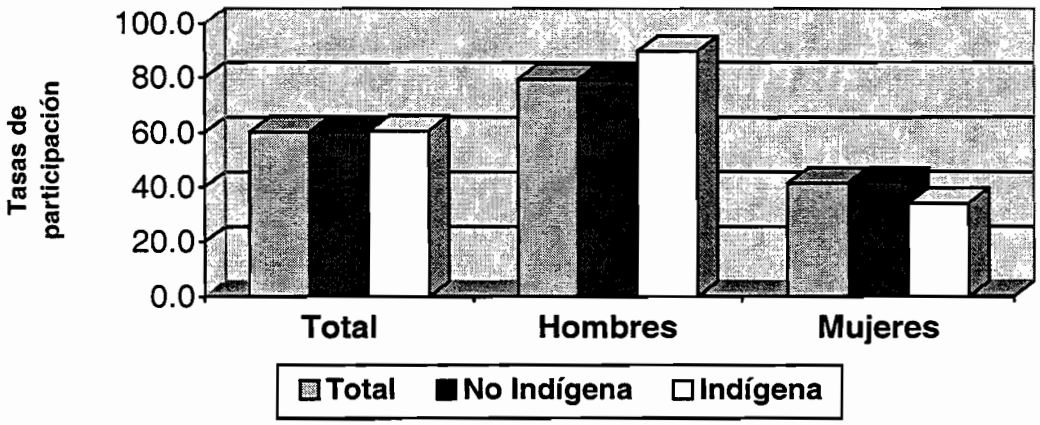
1V. Resultados de la Encuesta

La información preliminar de la Encuesta de Hogares realizada entre los meses de agosto y septiembre de 2001 indica que la población total de 15 años y más de edad, en edad de trabajar en el país asciende a 1,952,693 personas, representando la población no indígena (1854,729) el 94.9% y la población indígena (97,964) el 5.1%.

Al relacionar la población económicamente activa con el total de la población en edad de trabajar se obtienen tasas de participación a nivel de la población total, no indígena e indígena de 60.3%, 60.3% y 60.6% respectivamente.

A nivel de áreas, se observa que 1,281,595 personas de 15 años y más de edad (65.6%) se ubican en comunidades urbanas, de este grupo 787,340 personas 61.4% participan activamente en el mercado laboral; en lo que se refiere al área rural se registran 671,098 personas de 15 años y más de edad de las cuales 390,592 personas participan de la actividad económica (58.2%).

TASAS DE PARTICIPACIÓN DE LA POBLACIÓN TOTAL, NO INDÍGENA E INDÍGENA: AGOSTO 2001



El análisis por sexo revela una mayor participación de los hombres en la actividad económica en relación las mujeres ya que mientras el 79.4% de la población masculina total de 15 años y más participa en la actividad económica sólo el 41.4% de la población femenina total está inserta en el mercado laboral.

Igual situación se observa en las tasas de participación por sexo de la población no indígena (78.9 % hombres, 41.8 % mujeres) e indígena (89.7% hombres y 34.1%).

Antes de presentar algunas consideraciones sobre la evolución del mercado laboral en los últimos años, es preciso considerar las modificaciones metodológicas introducidas en la Encuesta; en ese sentido, y con el objeto de mostrar algunos indicios en lo relativo a este tema, nos remitiremos a los resultados de la encuesta referidos a la población no indígena, aunque a diferencia del resto la del 2001 incluye a la población que reside en áreas dispersas.

Siendo así, se puede señalar que la población no indígena experimenta en relación al año 2000 un incremento en la población en edad de trabajar de 49,100 lo cual evidencia el impacto del crecimiento demográfico sobre el mercado laboral.

Por otro lado, al comparar la evolución de la población económicamente activa de agosto 2000 a la fecha, se observa un ligero incremento en la tasa de participación (59.9% y 60.3%) respectivamente, lo que podría explicarse por la incorporación por primera vez de las áreas rurales de difícil acceso en este tipo de encuestas.

Las mayores tasas de participación en la actividad económica se registran en las provincias de Darién (64.0%), Veraguas (62.5% y Panamá (62.4%) y las tasas de participación más bajas en las provincias de Bocas del Toro (51.0%), Chiriquí (54.8%) y Coclé (55.2%).

TASAS DE PARTICIPACIÓN Y DESOCUPACIÓN EN LA REPÚBLICA: AGOSTO 20001

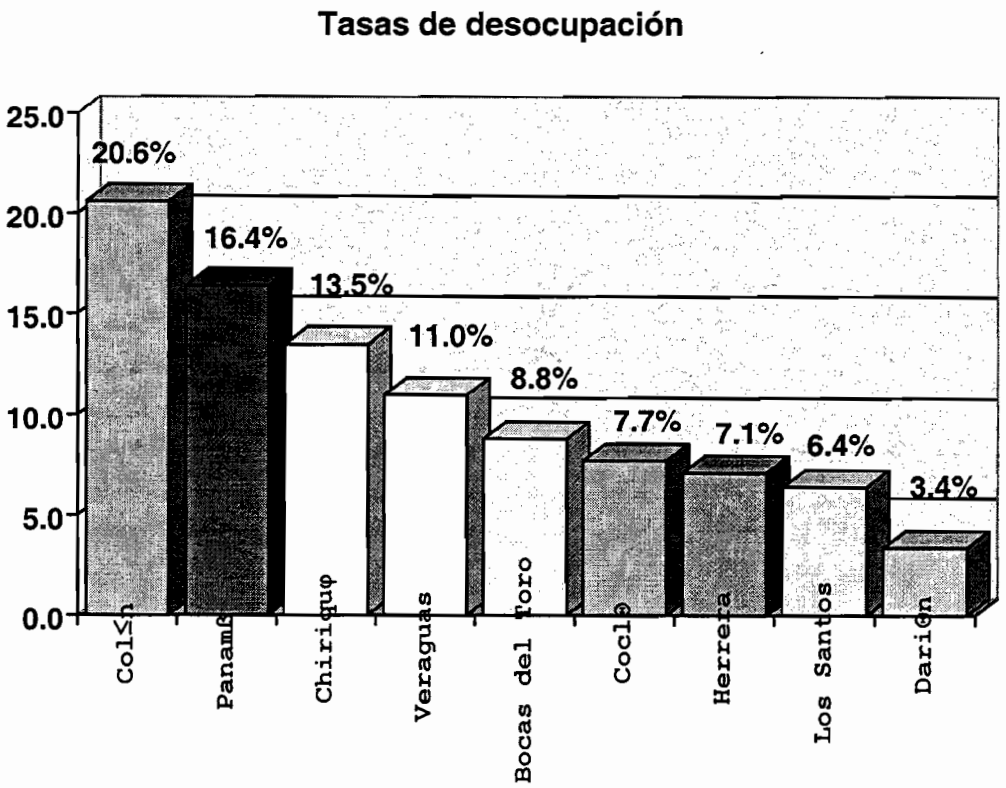
Desagregación geográfica	Tasas de participación por cien habitantes	Tasas de desocupación
Total	60.3	13.7
No indígena	60.3	14.4
Bocas del Toro	51.0	8.8
Coclé	55.2	7.7
Colón	59.9	20.6
Chiriquí	54.8	13.5
Darién	64.0	3.4
Herrera	60.8	7.1
Los Santos	59.6	6.4
Panamá	62.4	16.4
Distrito de Panamá	64.4	17.1
Ciudad de Panamá	64.6	14.8
Resto del Distrito de Panamá	64.1	20.7
Distrito de San Miguelito	62.8	18.6
Panamá Oeste	58.4	13.2
Resto de la Provincia de Panamá	56.6	10.4
Veraguas	62.5	11.0
Area Indígenas	60.6	1.2

En lo relativo al nivel de desocupación se registra a nivel total del país una tasa de desocupación de 13.7%, a nivel de la población no indígena de 14.4% y a nivel de la población indígena de 1.2%.

Al comparar la tasa desocupación de la población no indígena de agosto 2000 a la fecha se observa un incremento de casi un punto en el porcentaje de desocupación ya que la tasa pasó de 13.5% a 14.4% durante el período en mención.

En lo que se refiere al comportamiento de este indicador por área se observa en el área urbana una tasa de desocupación de 16.6%, en contraste con un 7.8% de desocupación en el área rural.

Al analizar la tasa de desocupación por provincia se registran los mayores niveles de desocupación en las provincias de Colón (20.6%) y Panamá (16.4%); en contraposición los niveles más bajos se observan en el área indígena (1.2%) y en las provincias de Darién (3.4%), Los Santos (6.4%) y Herrera (7.1%).



En lo referente a la inserción de la población no indígena en la economía, la Encuesta revela que el 19% de la población ocupada está en la actividad de Comercio al por mayor y al por menor, reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos; y el 17% en actividades relacionadas con la agricultura, ganadería, caza y silvicultura. Por otro lado, la encuesta revela a su vez que el mayor porcentaje de población desocupada (18%) trabajaba anteriormente en actividades relacionadas con el Comercio al por mayor y al por menor, reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos.

Respecto a las ocupaciones declaradas por la población ocupada el 17% trabaja como vendedores ambulantes, en servicios no clasificados en otros grupos, como obreros y jornaleros, etc.; el 14% son trabajadores de los servicios y vendedores de comercio y mercados y el 11% son empleados de oficina.

Al respecto de la población que declaró alguna ocupación el 63% indico ser empleada, el 30% es cuenta propia, el 4% es trabajador familiar y el 2% es patrono. De 639,628 personas registradas como empleadas, el 64% corresponde a la empresa privada, el 27.5% al gobierno y el 8.5% son empleados como servicios domésticos.

DISEÑO Y CONSTRUCCIÓN DE MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES

INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA (INEI)
PERÚ

ÍNDICE

Página

A. Los Censos de Población y Vivienda y los Marcos de Muestreo utilizados en las Encuestas de Hogares 265
1. Relación entre la cartografía censal y el marco de muestreo utilizado para la selección de las viviendas 265
2. Utilización del material censal para la construcción de los marcos de muestreo 265
3. Preparación y adaptación de los mapas censales para la elaboración del marco de muestreo 268
B. Diseño y Construcción del Marco de Muestreo 268
1. Importancia de los marcos de áreas para la selección de viviendas 268
2. Requerimientos de información, dominios de estudio e inferencia estadística 269
3. Los marcos maestros de muestreo y los diseños de muestras maestras 270
4. Procedimientos para la información de las unidades de marco de muestreo 274
5. Tamaño de las unidades primarias de muestreo y se efecto en la varianza de los estimadores 274
6. Tamaño óptimo de las unidades secundarias y terciarias de muestreo (segmentos o conglomerados de viviendas) y su influencia en la varianza de los estimadores 274
C. Aspectos Asociados a la Construcción del Marco de Muestreo 275
1. Procedimiento para la estratificación del marco de muestreo 275
2. Selección de unidades de muestreo y requerimientos de información 276
3. Métodos para la asignación de la muestra 279
4. Procedimientos para la selección de la muestra 283
5. Diseño en tres etapas: ventajas y desventajas 283
6. Influencia del nuevo marco muestral sobre los resultados de la encuesta 283
7. Muestras de panel de hogares y viviendas 284
8. Uso de proyecciones de población para generar estimaciones 287
D. Aspectos Operativos para la Construcción del Marco Muestreo 287
1. Creación de áreas de listado (diferencias entre áreas urbanas y rurales) 287
2. Procedimientos operativos para la elaboración de listados de viviendas 288
3. Actualización permanente del marco de muestreo y de la muestra 288
4. Tratamiento de viviendas, adicionales, omitidas y desaparecidas 289
5. Tratamientos de estratos especiales de vivienda 289

A. Los Censos de Población y Vivienda y los Marcos de Muestreo Utilizados en las Encuestas de Hogares

1. Relación entre la cartografía censal y el marco de muestreo utilizado para la selección de las viviendas

El último Censo Nacional de Población y Vivienda fue realizado en 1993. El Instituto Nacional de Estadística e Informática (INEI) en base a la cartografía y los resultados censales elaboró un marco muestral de áreas el cual ha tenido vigencia hasta el 2000.

En 1999, como tarea preparatoria de los censos de la década del 2000, el INEI realizó una actualización de la cartografía censal de 1993 y un registro exhaustivo de viviendas en los centros poblados urbanos del país. A partir de esa información se ha elaborado un nuevo marco muestral de áreas el cual está vigente para la selección de las viviendas de las investigaciones que se vienen realizando desde el 2001.

La cartografía censal está organizada a nivel de distritos, última unidad política administrativa del país. Es decir, para cada distrito se cuenta con 2 documentos cartográficos: el plano urbano y el mapa distrital. Estos documentos son utilizados como herramienta auxiliar en la Encuesta Nacional de Hogares (ENAHO) y otras investigaciones que realiza el INEI tales como la Encuesta Demográfica y Salud Familiar (ENDES) de 1996 y 2000.

El plano urbano del distrito es utilizado como material auxiliar para la identificación de la muestra de viviendas en los conglomerados urbanos seleccionados, mientras que el mapa distrital lo es para los conglomerados rurales seleccionados.

2. Utilización del material censal para la construcción de los marcos de muestreo

Hasta 1990, el marco muestral para las encuestas efectuadas por el INEI estaba basado en la cartografía y datos provenientes del Censo de Población y Vivienda de 1981 con una actualización parcial en las grandes ciudades seleccionadas para cada investigación en particular. No se contaba con marcos de muestreo en archivos magnéticos.

En 1991, el INEI actualizó cifras con el Pre-Censo de 1991, realizando una actualización de la cartografía en todos los distritos del país y levantando un registro de unidades de viviendas en todos los centros poblados urbanos. Esta información sirvió como marco muestral para las encuestas de hogares tales como la Encuesta Demográfica y de Salud Familiar de 1991 (ENDES II), Encuesta de Seguimiento del Consumo (ENSECO) del mismo año y, la Encuesta Nacional de Presupuestos Familiares (ENAPROM) de 1993 - 1994.

En 1993, el INEI ejecutó los Censos Nacionales: IX de Población y IV de Vivienda. Sobre la base de la cartografía y los resultados se elaboró un marco muestral tomando como base las unidades de área del censo. Este marco, básicamente de áreas, estaba conformado por un marco estratificado (según tamaño de viviendas) de las unidades primarias de muestreo (UPM) y un marco estratificado (según características socioeconómicas) de las unidades secundarias de muestreo (USM). Estos, fueron almacenados en medios magnéticos. La información del marco de USM fue actualizada en 1996 en las ciudades más importantes del país seleccionadas para la Encuesta Nacional Demográfica y de Salud Familiar (ENDES III) de ese año. Las unidades de muestreo, que conformaban el marco muestral eran:

UPM:

- **Unidad Primaria de Muestreo Urbana**
  Centro poblado o aglomerado con 2 mil y más habitantes.
- **Unidad Primaria de Muestreo Rural**
  Centro poblado o localidad con menos de 2 mil habitantes, para lo cual se definieron dos tipos de UPM's rurales:
  i) La Sección Censal Rural, con aproximadamente 4 Áreas de Empadronamiento Rural

ii) El centro poblado con 500 a menos de 2 mil habitantes.

USM :

- **Unidad Secundaria de Muestreo Urbana**

Conglomerado (sección censal urbana), conformado por una o más manzanas contiguas que en conjunto tienen aproximadamente 150 viviendas particulares.

- **Unidad Secundaria de Muestreo Rural**

Se consideran dos tipos de USM rural:

i) El Área de Empadronamiento Rural (AER) del censo, conformado por uno más centros poblados que en conjunto suman aproximadamente 100 viviendas particulares.

ii) El conglomerado (sección censal urbana) de 150 viviendas particulares en promedio.

UTM:

- **Unidad Terciaria de Muestreo**

Conformadas por las viviendas particulares dentro de los conglomerados.

En el cuadro 1, se observa la estructura del marco muestral utilizado para las ENAHO.

CUADRO 1. DISTRIBUCION DEL NUMERO DE UPM, USM Y UTM EN LA POBLACION, SEGUN DEPARTAMENTOS- CENSOS DE 1993

DEPARTAMENTO	POBLACION (CPV 1993)							
	UPM					USM		UTM
	CENTROS POBLADOS 2000 Y MAS HABITANTES			CCPP CON 500 A MENOS 2000 HAB.	SECCION CENSAL RURAL	SECCION CENSAL URBANA	AER	VIVIENDAS
	TOTAL	AUTO	NO AUTO					
TOTAL	**503**	**50**	**453**	**1084**	**3766**	**29940**	**15508**	**5099592**
AMAZONAS	12	1	11	40	111	269	410	77584
ANCASH	22	2	20	98	258	1191	1058	244799
APURIMAC	10	1	9	55	163	327	729	105055
AREQUIPA	30	2	28	45	93	1711	357	216371
AYACUCHO	13	1	12	75	215	603	911	154540
CAJAMARCA	20	2	18	67	447	710	1743	296681
CUSCO	22	2	20	71	339	995	1422	247200
HUANCAVELICA	6	1	5	60	172	298	720	114274
HUANUCO	10	2	8	51	205	474	896	149808
ICA	25	4	21	22	60	852	252	135372
JUNIN	49	3	46	81	190	1391	756	249248
LA LIBERTAD	39	2	37	45	192	1578	801	285692
LAMBAYEQUE	31	3	28	18	89	1191	356	187185
LIMA 1/	35	4	31	118	140	12731	545	1530326
LORETO	20	2	18	8	123	578	528	118489
MADRE DE DIOS	2	1	1	5	11	85	37	14992
MOQUEGUA	4	2	2	15	23	284	88	38649
PASCO	13	1	12	25	51	309	209	55427
PIURA	61	7	54	78	183	1812	685	291748
PUNO	26	2	24	56	524	950	2236	321949
SAN MARTÍN	33	2	31	30	107	621	477	118090
TACNA	3	1	2	14	17	426	75	54583
TUMBES	11	1	10	6	10	239	38	33500
UCAYALI	6	1	5	1	43	315	179	58030

1/: Incluye a la Provincia Constitucional del Callao

El marco muestral actualmente vigente desde el 2001, se basa en la cartografía e información del precenso de 1999. Las unidades de muestreo de este marco están definidas de manera similar al marco de 1993. ***Cada centro poblado urbano*** está dividido en zonas censales, las cuales tienen en promedio 50 manzanas cada una. Además, cada zona censal está conformada por 5 o 6 secciones censales, cada una con un promedio de 80 viviendas particulares. Estas secciones censales pasan a conformar el marco muestral denominándose conglomerados urbanos. ***Cada centro poblado rural con 500 a menos de 2000 habitantes*** tiene la misma configuración que un centro poblado urbano, es decir, está dividido en zonas censales y éstos a su vez en secciones que tienen en promedio 80 viviendas particulares. Estas secciones pasan a conformar el marco muestral denominado conglomerados rurales.

Aquellos centros poblados rurales con menos de 500 habitantes, se agrupan en áreas de empadronamiento rural (AER), las cuales tienen en promedio 100 viviendas. Estas AER pasan a conformar el marco muestral denominándose conglomerados rurales.

Las unidades primarias de muestreo (UPM) del marco muestral se han agrupado en 8 estratos teniendo en cuenta el total de viviendas:

Área Urbana:	**Código**
1. Ciudades con 100 mil y más viviendas:	111
2. Ciudades con 20 mil y menos de 100 mil viviendas:	121
3. Ciudades con 10 mil a menos de 20 mil viviendas:	132
4. Ciudades con 4 mil a menos de 10 mil viviendas:	142
5. Centros poblados con 400 a menos de 4 mil viviendas:	201
Area Rural:	
6. Centros poblados con menos de 400 viviendas:	202
7. AER Simples con menos de 140 viviendas:	321
8. AER Compuestos con mas de 140 viviendas:	311

En el cuadro 2, se observa la estructura del marco muestral actual utilizado para las ENAHO a partir del 2001.

CUADRO 2. DISTRIBUCION DEL NUMERO DE UPM EN LA POBLACION, SEGUN DEPARTAMENTOS- PRECENSO DE 1999

Depto.	Total De Distritos	URBANO						RURAL		Total	
		Ciudades				CCPP en dist. Estándar		Áreas de Empadron. Rural (AER)		UPM	%
		100 mil y más viv. (500 mil y más hab.)	20 mil a 100 mil viv (100 mil a menos de 500 mil hab.)	10 mil a 20 mil viv. (50 mil a menos de 100 mil hab.)	4 mil a 10 mil viv. (20 mil a menos de 50 mil hab.)	400 a 4 mil viv. (2mil a menos de 20 mil hab)	Menos de 400 viv. (500 a menos de 2 mil hab.)	Com-pues-to	Sim-ple		
		(111)	(121)	(132)	(142)	(201)	(202)	(311)	(321)		
Amazonas	83				3	16	64	159	152	**394**	3,5
Ancash	166		2		1	38	122	429	292	**884**	7,8
Apurímac	80			1	1	11	64	250	110	**437**	3,8
Arequipa	109	1			3	31	58	115	128	**336**	3,0
Ayacucho	111		1		1	41	65	294	213	**615**	5,4
Cajamarca	127		1	1		27	97	811	447	**1.384**	12,2
Cusco	108		1		3	54	46	513	258	**875**	7,7
Huancaveli.	94				1	25	67	290	160	**543**	4,8
Huanuco	75		1		1	20	51	381	172	**626**	5,5
Ica	43		3		3	16	13	61	81	**177**	1,6
Junín	123		1		3	66	51	333	161	**615**	5,4
La Libertad	83	1		1	10	26	41	354	146	**579**	5,1
Lambayeq.	38		1		7	24	4	155	57	**248**	2,2
Lima	177	1	1	3	3	40	72	146	196	**462**	4,1
Loreto	49		1		1	15	29	212	75	**333**	2,9
M. de Dios	11				1	3	7	12	27	**50**	0,4
Moquegua	20			2		7	9	17	44	**79**	0,7
Pasco	28			1		18	7	110	34	**170**	1,5
Piura	64		2	4	6	32	18	303	222	**587**	5,2
Puno	108		2		3	44	59	710	484	**1.302**	11,4
San Martín	77			1	4	32	38	185	147	**407**	3,6
Tacna	26		1			4	18	21	29	**73**	0,6
Tumbes	12		1			9	2	14	22	**48**	0,4
Ucayali	14		1			7	5	89	59	**161**	1,4
Total	**1.826**	**3**	**20**	**14**	**55**	**606**	**1007**	**5964**	**3716**	**11.385**	**100,0**

3. Preparación y adaptación de los mapas censales para la elaboración del marco de muestreo.
Es importante mencionar que el marco de 1993 estaba computarizado, pero la cartografía básica no estaba digitalizada. En el marco muestral actual la digitalización de la cartografía está en proceso dándose prioridad a las ciudades más importantes del país.

B. Diseño y Construcción del Marco de Muestreo

1. Importancia de los marcos de áreas para la selección de viviendas

El marco muestral actual de la ENAHO está organizado en conglomerados y conformada por aproximadamente 55 mil a nivel nacional. Los conglomerados son un conjunto de áreas de empadronamiento que constan en promedio de 80 viviendas en el área urbana y de 100 viviendas en el

área rural. Este marco muestral está organizado de una manera computarizada para poder realizar la selección de la muestra de conglomerados de una manera rápida y con la aleatoriedad que se requiere. Una vez realizada la selección de los conglomerados se procede a acceder a la base de datos del directorio de viviendas a fin de proceder a la selección automática de un grupo de viviendas del total contenido en el conglomerados seleccionado.

Los marcos muestrales tienen el problema de su rápida desactualización, esto ocurre con mayor frecuencia en las áreas periféricas de las grandes ciudades. Por esta razón la actualización del directorio de viviendas de cada conglomerados seleccionado antes de la ejecución de la encuesta es una tarea permanente en la ENAHO. Este procedimiento permite asegurar la probabilidad de selección de las nuevas viviendas, así como para determinar y ajustar los pesos y ponderaciones a nivel de cada conglomerado seleccionado.

2. Requerimientos de información, dominios de estudio e inferencia estadística

En la Encuesta Nacional de Hogares (ENAHO), estimaciones confiables son requeridas tanto para el nivel nacional como para niveles menores de desagregación como son las regiones naturales (8) o departamentos (24). Evidentemente, el tamaño de estas áreas varían considerablemente en términos de viviendas y habitantes. Así por ejemplo, la ciudad de Lima Metropolitana, por sí constituye un dominio de estudio por su tamaño e importancia, representa casi el 40% de la población total del país mientras que la región selva representa solo el 12%. También se tiene al departamento Madre de Dios que representa menos del 0.5% de la población del país.

Dado que la ENAHO es ejecutada en forma trimestral, los requerimientos de información varían por trimestre. Así, en el primer trimestre de encuesta el módulo principal se refiere a características de la vivienda y del hogar. En el segundo trimestre, se refiere a características de educación y salud. En el tercer trimestre, se releva información principalmente relacionada a características del empleo y en el cuarto trimestre, el módulo principal es niveles de vida y pobreza.

Por razones de costos, los dominios de estudio considerados en la ENAHO regularmente son las regiones naturales (costa, sierra y selva) con una muestra global de 8 mil viviendas particulares por trimestre. Excepcionalmente, en 1995, 1996 y el 2001 los dominios de estudio para el cuarto trimestre fueron los 24 departamentos con una muestra global cercana a las 20 mil viviendas particulares por cada trimestre.

El periodo de ejecución de la operación de campo por trimestres es la siguiente:

- 1er. Trimestre : Febrero - Marzo
- 2do. Trimestre : Mayo - Junio
- 3er. Trimestre : Agosto - Setiembre
- 4to. Trimestre : Noviembre - Diciembre

La distribución de la muestra es tanto para el área urbana y rural del país. En el siguiente cuadro se muestra según año y trimestre la distribución de la muestra:

TAMAÑO DE LA MUESTRA DE LA ENAHO, SEGÚN AÑO Y TRIMESTRE

AÑO	TRIMESTRE	Nº DE VIVIENDAS
1995	I - IV	20000
1996	I - IV	20000
1997	I	13000
	II	14000
	III	13000
	IV	7200
1998	I	2600 (*)
	II	7186
	III	9007
	IV	7923
1999	I	7190
	II	7626
	III	9070
	IV	4008
2000	I	1310 (*)
	II	2546
	III	11960
	IV	4359
2001	III	10 395
	IV	18 180

(*) = Sólo Lima Metropolitana.

Una de las limitaciones estadísticas de la información que existen en la Encuesta nacional de Hogares es el problema de la inferencia, es decir, debido al tamaño adoptado de la muestra en algunos trimestres no se puede efectuar inferencias a nivel departamental, provincial y distrital ya que estas no son representativas, en la mayoría de los casos el tamaño de la muestra obedece a factores de costo o presupuesto.

Por otro lado, la muestra panel reduce los niveles de inferencia, esto debido a que cada trimestre que se aplica la panel con respecto al año anterior se producen rechazos, lo cual reduce su tamaño a todas aquellas viviendas que tuvieron como resultado final en la encuesta completo e incompleto; esta situación repercute en los niveles de inferencia que ya no representan el 100 % de su diseño original.

3. Los marcos maestros de muestreo y los diseños de muestras maestras

En 1998, el INEI estableció la necesidad de elaborar una muestra maestra de áreas para el desarrollo de encuestas de propósitos múltiples (ENAHO), con el fin de hacer el seguimiento de un conjunto de variables que cubren diversos temas relacionados con el empleo y las condiciones de vida. La muestra maestra ha sido básicamente una muestra de áreas de empadronamiento censales agrupadas, y denominadas conglomerados; su propósito fue seleccionar diferentes submuestras para la implementación de encuestas más reducidas en alcance y cobertura durante el período en mención.

La muestra fue diseñada para obtener indicadores relacionados con: la composición del hogar, características de la vivienda, acceso a servicios, acceso a la educación, acceso a la salud, acceso a

programas sociales, empleo, ingreso, gasto y otros temas libres.

Para el diseño de la muestra se consideraron 6 variables:

i. Ingreso Per Capita por Persona
ii. Porcentaje de Personas que Acceden a Servicios de Salud
iii. Porcentaje de Hogares con Servicio de Desagüe
iv. Porcentaje de Hogares Beneficiados con Programas Sociales
v. Tasa de Desempleo Urbana

Se consideraron básicamente ocho (8) dominios de estudio,

i. Lima Metropolitana,
ii. Costa Norte,
iii. Costa Centro,
iv. Costa Sur.
v. Sierra Norte,
vi. Sierra Centro,
vii. Sierra Sur, y
viii. Selva Total.

También se consideró importante la desagregación en cinco niveles de urbanización de las unidades primarias de muestreo (UPM) agrupados en las siguientes categorías:

i. Menos de 500 habitantes,
ii. De 500 a 2 mil habitantes,
iii. De 2 mil a 20 mil habitantes,
iv. De 20 mil a 100 mil habitantes, y
v. Más de 100 mil habitantes

Para determinar el tamaño global de la muestra maestra se tomó en cuenta la variabilidad de las características o variables de estudio, tanto en las unidades de muestreo de segunda (USM) y tercera etapa (UTM).

Como se sabe, el error muestral para una encuesta del tipo de la ENAHO tiene dos componentes: el primero corresponde a la variación **entre** conglomerados y el segundo a la variación de las viviendas o los hogares **dentro** de los conglomerados.

La variabilidad **entre** conglomerados es mucho mayor para las áreas urbanas que en las áreas rurales, pero la variabilidad **dentro** del conglomerado seleccionado es menor en lo urbano que en lo rural.

Por lo tanto, un importante factor para controlar el error muestral es reducir el mayor componente de variabilidad que es la variación entre conglomerados, el cual es inversamente proporcional al número de conglomerados (especialmente en el área urbana).

Como consecuencia, para cumplir con los diferentes objetivos de la encuesta se consideró que un tamaño de muestra de aproximadamente 4 mil conglomerados era el tamaño de muestra más adecuado para producir con suficiente precisión los principales indicadores para los diferentes niveles de desagregación requeridos en la Encuesta Nacional de Hogares.

En la Encuesta Nacional de Hogares, se requieren estimaciones confiables tanto para el nivel nacional como para niveles menores de desagregación como son las regiones naturales o departamentos. Cuando estas áreas varían considerablemente en sus tamaños totales (habitantes o viviendas), surgen problemas en el uso de asignaciones estándar.

Por un lado se tiene la asignación de la muestra total del país entre sus dominios que *minimiza el error relativo al nivel nacional* , la cual es casi igual a la asignación proporcional al tamaño del dominio, ocasionando que ciertos dominios tengan un *mayor error relativo* que el tamaño de muestra producido por la directa minimización del error dentro del dominio.

Por otro lado, cuando se requiere que cada dominio tenga el *mismo nivel de error relativo* entonces puede resultar que el error relativo producido para el nivel nacional sea *mucho más grande* que el calculado con la respectiva minimización al nivel nacional. Esta asignación con el mismo nivel de error relativo produce el *mismo tamaño de muestra* para cada dominio sin consideración del tamaño o de la importancia, es decir el tamaño de muestra en el dominio es independiente de su tamaño.

Aprovechando el concepto de la importancia (tamaño) del dominio anteriormente discutido y tomando en consideración el concepto de precisión del indicador estadístico en cada dominio, entonces una mejor asignación de la muestra total entre sus dominios componentes se obtiene cuando el tamaño de muestra en el dominio es **función de la importancia del dominio y de su correspondiente error relativo** (véase Power Allocations Determining Sample Sizes for Subnational Areas por Michael D. Bankier. The American Statistician. August 1988 Vol 42 No 3).

Teniéndose el indicador estadístico Y (promedio, proporción o porcentaje) estimado en cada dominio, y además teniéndose la importancia X (tamaño) de cada dominio entonces se considera la minimización de la siguiente función en términos de la mejor asignación de muestra entre los dominios (n_h):

$$F = \sum_h (X_h^{\alpha} CV(Y_h))^2$$

con la restricción de que:

$$n = \sum_h n_h$$

donde :

$CV(Y_h)$: es el coeficiente de variación (error relativo) del indicador Y en el dominio h,

X_h^{α} : es la potencia α de la importancia (tamaño) X en el dominio h,

n_h : es la muestra asignada al dominio h,

α : es un valor constante entre 0 y 1.

Se demuestra matemáticamente que la función F es minimizada cuando la asignación por dominio es dada por la siguiente relación:

$$n_h = n \frac{S_h X_h^{\alpha} / Y_h}{\sum_h S_h X_h^{\alpha} / Y_h}$$

donde :

Sh : esta definido dentro de la fórmula de la varianza del indicador,

$$Var\ (Y_h) = S_h^2 / n_h\ ,\ \text{en el dominio h.}$$

El correspondiente ejercicio de este procedimiento fue llevado a cabo con varios indicadores requeridos y seleccionados en el estudio de la ENAHO 98, y para los cuales se tenía información de ellos con sus errores relativos de encuestas similares pasadas. Dichos indicadores fueron los siguientes :

i. Ingreso per capita por persona
ii. Porcentaje de personas que acceden a servicios de salud
iii. Porcentaje de hogares con servicio de desagüe
iv. Porcentaje de hogares que se benefician con programas sociales
v. Tasa de desempleo

Se realizó la asignación para cada dominio de cada uno de esos indicadores considerando el tamaño de muestra total de 3 mil 884 conglomerados.

CUADRO 3. DISTRIBUCION DE LA MUESTRA MAESTRA DE CONGLOMERADOS, POR AREA URBANA Y RURAL, SEGÚN DOMINIOS DE ESTUDIO-ENCUESTA NACIONAL DE HOGARES

DOMINIOS DE ESTUDIO	MUESTRA MAESTRA DE CONGLOMERADOS		
	TOTAL	URBANO	RURAL
TOTAL	**3884**	**2209**	**1675**
Costa	**893**	**702**	**191**
Costa norte	397	299	98
Costa centro	301	236	65
Costa Sur	195	167	28
Sierra	**1922**	**738**	**1184**
Sierra norte	590	211	379
Sierra centro	764	272	492
Sierra sur	568	255	313
Selva	**707**	**407**	**300**
Lima Metro.	**362**	**362**	**-**

Ventajas de la Muestra Maestra

Las principales ventajas obtenidas en la utilización de la muestra de propósitos múltiples han sido los siguientes:

- Reducción de costos asociados al trabajo de elaboración de los marcos de muestreo y la selección de muestras. Esto es importante dado que los costos y tiempos empleados en la elaboración de marcos de muestreo y selección de muestras, independientemente para cada trimestre o ronda de la encuesta, eran relativamente mayores a lo actual.
- Selección con mayor rapidez y economía las muestras trimestrales de ENAHO, utilizando mismos materiales de muestreo. Así mismo, el material cartográfico desde planos urbanos, mapas distritales y croquis, fue utilizado para varias rondas o trimestres de la encuesta.
- Ha permitido la selección de submuestras para diferentes tipos de encuestas.
- Ha facilitado la integración de información estadística recopilada en los diferentes trimestres de la ENAHO, al utilizar unidades de muestreo superpuestas.

Desventajas de la Muestra Maestra

Las principales desventajas en la utilización de la muestra de propósitos múltiples han sido las siguientes:

- La muestra maestra implementada imposibilita diseñar encuestas especiales para áreas geográficas menores, por ejemplo, provincias, áreas periféricas donde habita población de escasos recursos, grupos étnicos, etc..
- Al reutilizarse las unidades de la muestra, especialmente a nivel de hogares, cabe la posibilidad de producirse sesgos por efecto del condicionamiento.

4. Procedimiento para la formación de las unidades de marco de muestreo

Las unidades del marco muestral de áreas de la ENAHO son definidas teniendo en cuenta las áreas de empadronamiento conformadas para el empadronamiento en los censos nacionales de población y vivienda.

A partir de la segmentación del empadronamiento del Censo de Población y Vivienda de 1993, la sección censal urbana fue definida como la unidad secundaria de muestreo urbana. En el área rural, la sección censal rural y el área de empadronamiento rural (AER) fueron definidas como unidad primaria de muestreo y unidad secundaria de muestreo respectivamente.

Así mismo, la segmentación automática de las áreas de empadronamiento realizada a partir de información del Precenso de 1999, ha servido para definir las unidades de muestreo del nuevo marco muestral actualmente vigente para la ENAHO. La sección censal urbana ha sido nuevamente definida como la unidad secundaria de muestreo urbana. En el área rural, el área de empadronamiento rural (AER) ha sido definida como unidad primaria de muestreo.

5. Tamaño de las unidades primarias de muestreo y su efecto en la varianza de los estimadores

Como ya se ha mencionado antes, las unidades primarias de muestreo (UPM) contenidas en el marco muestral actual de la ENAHO están agrupadas en 8 estratos según el tamaño de viviendas. Esta estratificación tiene como objetivo minimizar los errores muestrales en la primera etapa de selección. Evidentemente, los tamaños altamente variables de estas unidades pueden conducir a una sobreestimación de la varianza de las estimaciones por lo que en la ENAHO para el cálculo de las varianzas se utiliza el método de los conglomerados últimos para la estimación de dichos errores de muestreo.

6. Tamaño optimo de las unidades secundarias y terciarias de muestreo (segmentos o conglomerados de viviendas) y su influencia en la varianza de los estimadores

En las encuestas de hogares tales como la ENAHO, el conglomerado es normalmente una manzana o varias manzanas contiguas, un centro poblado, una aldea o caserío, etc. Dentro de estas unidades, naturalmente, se correlacionan las variables socioeconómicas entre los hogares. Esto contradice la meta teórica de tener un alto grado de heterogeneidad en el conglomerado. Como consecuencia, la alta correlación intraclase hace que aumente la varianza del estimador.

La experiencia práctica demuestra que para las encuestas de hogares los valores óptimos para el número de viviendas muestrales se encuentran en la vecindad de 8 a 16 viviendas por conglomerado. En la ENAHO, el tamaño de viviendas muestrales por conglomerado urbano es de 8 en el área urbana y 12 en el área rural.

C. Aspectos Estadísticos Asociados a la Construcción del Marco de Muestreo

1. Procedimientos para la estratificación del marco de muestreo

En toda encuesta de hogares en la que se haga uso del muestreo multietápico, la estratificación de las unidades de muestreo es fundamental e importante para la reducción de la varianza. A fin de mejorar la eficiencia del marco muestral de la Encuesta Nacional de Hogares (ENAHO) utilizado desde 1995, se tomaron criterios de estratificación para las unidades de muestreo.

En una primera etapa, las unidades primarias de muestreo (UPM) fueron agrupadas en estratos teniendo en cuenta su importancia en tamaño de habitantes:

- **Estrato 1**: Unidades primarias de muestreo (UPM) urbanas de gran tamaño, tales como ciudades con 100 mil y más habitantes.
- **Estrato 2**: Unidades primarias de muestreo urbanas de mediano tamaño, tales como ciudades con 20 mil a menos de 100 mil habitantes.
- **Estrato 3**: Unidades primarias de muestreo urbanas menores, tales como aquellos centros poblados o localidades con 2 mil a menos de 20 mil habitantes.
- **Estrato 4**: Unidades primarias de muestreo rurales concentradas, tales como centros poblados o localidades con 500 a menos de 2 mil habitantes.
- **Estrato 5**: Áreas rurales pequeñas y dispersas, formadas por centros poblados con menos de 500 habitantes, agrupados en AER (áreas de empadronamiento rural).

Es importante mencionar que los estratos 1,2 y 3, conforman el área urbana mientras que los estratos 4 y 5, el área rural. Así mismo, la ENAHO es comparable con el Censo considerando solamente el estrato 5 como área rural.

La inclusión de conglomerados o unidades secundarias de muestreo (USM) en la muestra, dado el carácter multitemático de la ENAHO, tiende a incrementar la varianza muestral y el método que se diseña para minimizar la varianza de una variable, pudiera conducir a resultados no satisfactorios para otra. En tal sentido, deben buscarse indicadores socio económicos que determinen diferencias que afecten a la mayor parte de las características.

La mayoría de las variables importantes analizadas en la ENAHO tales como empleo, ingresos, gastos, educación y salud, están altamente correlacionadas con los niveles socioeconómicos de la población. Por lo que fue necesario definir subestratos socioeconómicos en los aglomerados urbanos de gran tamaño, como son las grandes ciudades.

A fin de reducir la varianza de la segunda componente, la estratificación de las USM fue un objetivo primordial para la ENAHO.

Sobre la base de información de los censos de 1993, el INEI procesó y obtuvo 5 indicadores de Necesidades Básicas Insatisfechas (NBI) para los hogares peruanos, al nivel de distritos:

- Hogares en viviendas con características físicas inadecuadas
- Hogares en viviendas con hacinamiento
- Hogares en viviendas sin servicio de desagüe de ningún tipo
- Hogares con niños que no asisten a la escuela
- Hogares con alta dependencia económica

La medición estaba referida a los hogares que habitaban en viviendas particulares con ocupantes presentes. Haciendo uso de esta metodología, estos indicadores se procesaron para cada USM del marco de las UPM, según el procedimiento siguiente:

1. En primer lugar, se identifican a los hogares carentes por cada indicador; luego, para el cálculo de la proporción de hogares carentes por cada indicador (Ij), se le relaciona con el total de hogares de la USM correspondiente.

2. Se agrupan los hogares que tienen al menos 1 NBI, de la manera siguiente:

HOG1 : Hogares con 1 necesidad básica insatisfecha.

HOG2: Hogares con 2 necesidades básicas insatisfechas.

HOG3: Hogares con 3 necesidades básicas insatisfechas.

HOG4: Hogares con 4 necesidades básicas insatisfechas.

HOG5: Hogares con 5 necesidades básicas.

3. Basándose en estos cinco grupos de hogares se obtiene por sumatoria y agregación el índice de necesidades básicas insatisfechas (NBI), que representa el indicador sintético de los hogares con al menos una necesidad básica insatisfecha para cada USM, resumida en la siguiente expresión matemática:

$$INBI = \Sigma\ (HOGi\ /\ HT)\ ;$$

Donde:

INBI: Indice de Necesidades Básicas Insatisfechas para cada USM

HOGi: Hogares con al menos una necesidad básica insatisfecha

HT: Total de hogares en la USM

4. Se ordenaron las USM según este indicador para cada UPM. Con esto se obtiene una estratificación implícita del marco cuando la muestra es seleccionada sistemáticamente con probabilidad proporcional al tamaño.

El uso de la metodología del NBI, tuvo como propósito obtener muestras estratificadas que aseguren la representatividad de los diferentes niveles socioeconómicos de la ciudad investigada.

2. Selección de las unidades de muestreo y requerimientos de información

La muestra de la Encuesta Nacional de Hogares (ENAHO) se basa en un diseño probabilístico, estratificado y de conglomerados que considera básicamente 3 etapas de muestreo. La selección de las unidades de muestreo se realiza a partir de marcos de muestreo elaborados para cada etapa de muestreo.

En el área urbana, la muestra es trietápica, de selección sistemática proporcional al tamaño de viviendas en la primera y segunda etapa, y sistemática simple con arranque aleatorio en la tercera etapa. El número de unidades primarias de muestreo se determina en función del número de conglomerados requeridos en el área urbana para cada dominio de estudio. Se procede a seleccionar las UPM´s mediante el muestreo sistemático con probabilidad proporcional al tamaño en número de viviendas.

Previa a la selección, se realiza una estratificación implícita de las UPM´s teniendo en cuenta:

- Los niveles de estratificación por tamaño (estratos I, II, III, IV y V) y,
- La continuidad geográfica (norte a sur)

La simbología utilizada es la siguiente:

nh : Tamaño muestral de UPM´s en el área urbana del h-dominio.

Mh : Total de viviendas en el área urbana del h-dominio, según precenso 1999.

La fracción de muestreo en la primera etapa, o la probabilidad que la i-UPM, en el h- dominio de estudio, sea seleccionada en la muestra es:

$$f1 = nh * Zhi \quad ; \quad \text{donde:}$$

Zhi = Mhi / Mh : Probabilidad de seleccionar la i-UPM en el área urbana del h-dominio de estudio .
Mhi : Total de viviendas en la i-UPM, según el precenso de 1999.

La unidad de selección en la segunda etapa es el conglomerado (USM). La simbología utilizada es la siguiente:

ghi : Muestra de USM´s en la i-UPM urbana seleccionada del h- dominio.
Mhi : Total de viviendas en la i-UPM, según el precenso de 1999.

Se procede a seleccionar las USM en las UPM en forma sistemática con probabilidad proporcional al tamaño de viviendas particulares.

K2 = Mhi / ghi : Intervalo de selección de las USM.

La fracción de muestreo en la segunda etapa, o la probabilidad que la j-USM, en la i-UPM urbana del h-dominio sea seleccionada en la muestra es:

$$f2 = ghi * Zhij \quad , \quad \text{donde:}$$

Zhij = Mhij / Mhi : Probabilidad de seleccionar la j-USM, en la i-UPM urbana seleccionada en el h-dominio.

Mhij : Total de viviendas en la j-USM, en la i-UPM urbana seleccionada, en el h-dominio, según precenso de 1999.

La unidad de selección en la tercera etapa es la vivienda particular. En cada USM seleccionada, se procede a la selección de un número fijo de viviendas con arranque aleatorio y en forma sistemática simple al azar. La simbología es la siguiente:

K3 = M'hij / mhij donde:

M'hij : Total de viviendas en la j-USM, de la i-UPM urbana seleccionada del h-dominio, según registro actualizado de viviendas.

mhij : Tamaño muestral de UTM (viviendas) en la j-USM, en la i-UPM urbana del h-dominio.

La fracción de muestreo en la tercera etapa o la probabilidad que la k-UTM, en la j-USM, de la i-UPM, en el h-dominio de estudio sea seleccionada en la muestra es:

$$f3 = mhij * Zhjk \quad ; \quad \text{donde:}$$

Zhijk = 1 / M'hij : Probabilidad de seleccionar la k-UTM, en la j-USM, en la i-UPM, en el h-dominio.

En el área rural, la muestra es trietápica, de selección sistemática proporcional al tamaño de viviendas en la primera y segunda etapa, y sistemática simple con arranque aleatorio en la tercera etapa.

El número de Unidades Primarias de Muestreo se determina en función del número de conglomerados requeridos en el área rural para cada dominio de estudio.

Se procede a seleccionar las UPM´s mediante el muestreo sistemático con probabilidad proporcional al tamaño en número de viviendas.

Previa a la selección, se realiza una estratificación implícita de las UPM´s teniendo en cuenta:

- Los niveles de estratificación por tamaño (estratos VI, VII y VIII) y,
- La continuidad geográfica (norte a sur)

La simbología utilizada es la siguiente:

nh : Tamaño muestral de UPM´s en el área rural del h-dominio.

Mh : Total de viviendas en el área rural del h-dominio, según el precenso.

La fracción de muestreo en la primera etapa, o la probabilidad que la i-UPM rural, en el h-dominio de estudio, sea seleccionada en la muestra es:

$f1 = nh * Zhi$; donde:

$Zhi = Mhi / Mh$: Probabilidad de seleccionar la i-UPM, en el área rural del h-dominio.

Mhi : Total de viviendas en la i-UPM rural, según el precenso.

La unidad de selección en la segunda etapa es el conglomerado (USM). La simbología utilizada es la siguiente:

ghi : Muestra de USM´s en la i-UPM rural seleccionada del h-dominio.

Mhi : Total de viviendas en la i-UPM rural seleccionada, del h-dominio.

Se procede a seleccionar las USM en las UPM en forma sistemática con probabilidad proporcional al tamaño de viviendas particulares.

$K2 = Mhi / ghi$: Intervalo de selección de las USM.

La fracción de muestreo en la segunda etapa, o la probabilidad que la j-USM, en la i-UPM rural del h-dominio sea seleccionada en la muestra es:

$f2 = ghi * Zhij$, donde:

$Zhij = Mhij / Mhi$: Probabilidad de seleccionar la j-USM, en la i-UPM rural seleccionada en el h-dominio.

Mhij : Total de viviendas en la j-USM en la i-UPM rural , en el h-dominio.

La unidad de selección en la tercera etapa es la vivienda particular. En cada USM seleccionada, se procede a la selección de un número fijo de viviendas con arranque aleatorio y en forma sistemática simple al azar. La simbología a utilizarse es la siguiente:

$k3 = M'hij / mhij$ donde:

M'hij : Número de viviendas en la j-USM, de la i-UPM rural seleccionada del h-dominio, según registro actualizado de viviendas.

mhij : Tamaño muestral de UTM (viviendas) en la j-USM, en la i-UPM rural del h-dominio.

La fracción de muestreo en la tercera etapa o la probabilidad que la k-UTM, en la j-USM, de la i-UPM rural, en el h-dominio de estudio sea seleccionada en la muestra es:

f3 = mhij * Zhjk ; donde:

Zhijk = 1 / M'hij : Probabilidad de seleccionar la k-UTM, en la j-USM, en la i-UPM rural, en el h-dominio.

3. Métodos para la asignación de la muestra

En la Encuesta Nacional de Hogares (ENAHO), una vez culminado el ingreso de la información a una base de datos, se procede a calcular los errores de muestreo de las variables o características más importantes del trimestre o ronda de encuesta. La fórmula del error de muestreo tiene en cuenta la estratificación y conglomeración utilizadas en el diseño muestral. Para lo cual se ha empleando la metodología incorporada en CENVAR del software estadístico IMPSS.

El error de muestreo se mide con el error estándar, el cual se define como la raíz cuadrada de la varianza del estimador, y no es mas que una medida de su variación en todas las muestras, es decir, el error estándar mide el grado de precisión con que el estadístico obtenido a partir de la muestra se aproxima al resultado que se hubiese obtenido al entrevistar a toda la población.

Las estimaciones obtenidas de la muestra pueden tener una alta o poca confiabilidad estadística. Es decir, pueden tener una muy buena precisión o servir solamente como información referencial. El conocimiento de estos indicadores de variabilidad permiten evaluar los tamaños de muestra considerados en la encuesta y ayudan en la determinación de los tamaños de muestra y asignación muestral para las encuestas futuras.

En la Encuesta Nacional de Hogares de 1995 y 1996, estimaciones confiables eran requeridas tanto para el nivel nacional como para niveles menores de desagregación como son los departamentos y ciudades capitales. La asignación de la muestra de viviendas por departamento fue realizada utilizando el promedio de la asignación proporcional y la asignación uniforme o igual. El método usado para asignar la muestra dentro del departamento se basaba en una distribución proporcional en función del estrato y respetando la división de urbano y rural.

En el área urbana, excluyendo la capital departamental, a fin de incrementar la eficiencia de campo y la supervisión, se seleccionaron en promedio 4 USM en las UPM no autorepresentadas y múltiplo de 4 en las UPM autorepresentadas.

En el área rural, a fin de asegurar una buena dispersión de la muestra, se seleccionaron a lo más 2 USM por cada UPM seleccionada.

Se consideró 8 viviendas por USM en área urbana y 12 viviendas por USM en área rural.

CUADRO 4. DISTRIBUCION DE LA MUESTRA DE VIVIENDAS, POR AREA URBANA Y RURAL, SEGUN DEPARTAMENTO - ENCUESTA NACIONAL DE HOGARES 1995 – I

DEPARTAMENTO	TOTAL	URBANO	RURAL
TOTAL	**20224**	**14992**	**5232**
AMAZONAS	656	416	240
ANCASH	992	752	240
APURIMAC	580	376	204
AREQUIPA	720	552	168
AYACUCHO	704	488	216
CAJAMARCA	632	392	240
CUZCO	808	544	264
HUANCAVELICA	536	320	216
HUANUCO	600	384	216
ICA	932	728	204
JUNIN	928	664	264
LA LIBERTAD	916	688	228
LAMBAYEQUE	896	680	216
LIMA	2920	2680	240
LORETO	688	496	192
MADRE DE DIOS	520	328	192
MOQUEGUA	652	424	228
PASCO	624	384	240
PIURA	1376	1136	240
PUNO	896	632	264
SAN MARTIN	896	680	216
TACNA	596	416	180
TUMBES	584	416	168
UCAYALI	572	416	156

En las ENAHO de 1997 y 1998, se consideraron básicamente 8 dominios urbanos de estudio :

- Costa Norte
- Costa Centro
- Costa Sur
- Sierra Norte
- Sierra Centro
- Sierra Sur
- Selva
- Lima Metropolitana

La asignación de la muestra se realizó básicamente en función del nivel de inferencia dominio de estudio y utilizando los indicadores de variabilidad obtenidas para las variables más importantes de la ENAHO de 1995 y 1996.

Mientras que el método usado para repartir la muestra en las ENAHO de 1995,1996 y 1997 se basaba en una distribución de la muestra de conglomerados en forma proporcional en función al tamaño del dominio de estudio, en la metodología de la ENAHO a partir del cuarto trimestre de 1998, en que se diseña la muestra maestra de conglomerados, se propone el uso del método de **Asignación de Potencia.**

El problema inmediato que se encaraba era la asignación de la muestra total del país entre los dominios con un determinado procedimiento de asignación y de como hacerlo cuando los tamaños de esos dominios variaban considerablemente en sus tamaños totales (población o viviendas), ya que cada uno de estos tamaños miden la importancia del correspondiente dominio.

Por un lado se tiene la asignación de la muestra total del país entre sus dominios que *minimiza el error relativo al nivel nacional* , la cual es casi igual a la asignación proporcional al tamaño del dominio, ocasionando que ciertos dominios tengan un *mayor error relativo* que el tamaño de muestra producido por la directa minimización del error dentro del dominio.

Por otro lado, cuando se requiere que cada dominio tenga el *mismo nivel de error relativo* entonces puede resultar que el error relativo producido para el nivel nacional sea *mucho más grande* que el calculado con la respectiva minimización al nivel nacional. Esta asignación con el mismo nivel de error relativo produce el *mismo tamaño de muestra* para cada dominio sin consideración del tamaño o de la importancia, es decir el tamaño de muestra en el dominio es independiente de su tamaño.

Estos casos extremos de la importancia de los dominios puede ser ilustrado matemáticamente en el siguiente ejemplo: Considérese un país A que tiene tres dominios y que se desea asignar una muestra de 600 casos entre esos dominios

DOMINIOS	TAMAÑO (X)
TOTAL	**1 101 000**
1	1 000 000
2	100 000
3	1 000

ahora bien, denominando X al tamaño del dominio, se define la regla de la distribución de la muestra como la asignación proporcional a X^{α} en cada dominio, donde α es un valor entre 0 y 1.

En el cuadro siguiente se observa la asignación de la muestra implementada para cinco valores de α: 0, 0.2, 0.5, 0.8 y 1.

DOMI-NIOS	$\alpha = 0$		$\alpha = 0,2$		$\alpha = 0,5$		$\alpha = 0,8$		$\alpha = 1$	
	X^0	m	$X^{0,2}$	m	$X^{0,5}$	m	$X^{0,8}$	m	X^1	M
TOTAL	**3**	**600**	**29,9**	**600**	**1348**	**600**	**73347**	**600**	**1101000**	**600**
1	1	200	15,9	319	1000	445	63096	516	1000000	545
2	1	200	10,0	201	316	141	10000	82	100000	54
3	1	200	4,0	80	32	14	251	2	1000	1

Como se puede observar en el cuadro anterior se tiene en un extremo la **asignación proporcional al tamaño** del dominio (X^1) con la consiguiente asignación de la muestra de 545 casos en el dominio 1, y con esta asignación, la muestra en el dominio 3 (la más pequeña) es de 1 solo caso.

Obviamente una mejor asignación de muestra para este dominio debería ser buscada para un valor de α menor o igual a 0,5.

En el otro extremo, se tiene la **asignación de igual tamaño** de muestra (la cual es equivalente a ser proporcional a X^0) para cada dominio que da una tremenda importancia al dominio más pequeñ0 con un enorme sacrificio de la importancia el dominio más grande.

Obviamente la mejor asignación al dominio más grande debería ser buscada para valores de α mayor o igual a 0,5.

Una mejor asignación de la muestra con esta regla (proporcional a X^{α}) para cada dominio puede ser obtenida para α igual a 0.5, es decir una asignación de muestra a cada dominio proporcional a la raíz cuadrada de X.

Aprovechando el concepto de la importancia (tamaño) del dominio anteriormente discutido y tomando en consideración el concepto de precisión del indicador estadístico en cada dominio, entonces una mejor asignación de la muestra total entre sus dominios componentes se obtiene cuando el tamaño de muestra en el dominio es **función de la importancia del dominio y de su correspondiente error relativo** (vease Power Allocations Determining Sample Sizes for Subnational Areas por Michael D. Bankier. The American Statistician. August 1988 Vol 42 No 3).

Teniéndose el indicador estadístico Y (promedio, proporción o porcentaje) estimado en cada dominio, y ademas teniéndose la importancia X (tamaño) de cada dominio entonces se considera la minimización de la siguiente función en terminos de la mejor asignación de muestra entre los dominios (n_h):

$$F = \sum_h (X_h^{\alpha} CV(Y_h))^2$$

con la restricción de que:

$$n = \sum_h n_h$$

donde :

$CV(Y_h)$: es el coeficiente de variación (error relativo) del indicador Y en el dominio h,
X_h^{α} : es la potencia α de la importancia (tamaño) X en el dominio h,
n_h : es la muestra asignada al dominio h,
α : es un valor constante entre 0 y 1.

Se demuestra matemáticamente que la función F es minimizada cuando la asignación por dominio es dada por la siguiente relación:

$$n_h = n \frac{S_h X_h^{\alpha} / Y_h}{\sum_h S_h X_h^{\alpha} / Y_h}$$

donde :

Sh : esta definido dentro de la fórmula de la varianza del indicador,

$Var(Y_h) = S_h^2 / n_h$, en el dominio h.

El correspondiente ejercicio de este procedimiento fue llevado a cabo con varios indicadores requeridos y seleccionados en el estudio de la ENAHO 98, y para los cuales se tenía información de ellos con sus errores relativos de encuestas similares pasadas. Dichos indicadores fueron los siguientes :

i. Ingreso per capita por persona
ii. Porcentaje de personas que acceden a servicios de salud
iii. Porcentaje de hogares con servicio de desague
iv. Porcentaje de hogares que se benefician con programas sociales
v. Tasa de desempleo

4. Procedimientos para la selección de la muestra

Los insumos del Marco Muestral de Áreas (Base de Datos y la Cartografía Censal) de la Encuesta Nacional de Hogares (ENAHO), tienen una correspondencia buinívoca que permite la selección de la muestra unidades de muestreo (UPM: distritos, centros poblados, AER), unidades secundarias de muestreo (USM: conglomerados), y unidades últimas de muestreo (UTM: viviendas).

El marco muestral computarizado comprende principalmente las distintas unidades muestrales de selección. Estas a su vez, registran información en campos tales como: código de identificación única de la unidad de muestreo, variables de selección, características de la estratificación, indicadores estadísticos y económicos, códigos de ubicación geográficos, datos de control, etc.

5. Diseño en tres etapas: ventajas y desventajas

La muestra de la Encuesta Nacional de Hogares (ENAHO) se basa en un diseño probabilístico, estratificado y de conglomerados que considera básicamente 3 etapas de muestreo. En la primera etapa, se seleccionan ciudades, centros poblados y/o áreas de empadronamiento rural. Estas unidades de muestreo varían en tamaño por lo que se procede a agruparlos en estratos para minimizar su efecto en la varianza de las estimaciones. En una segunda etapa, se seleccionan conglomerados que son de tamaño homogéneo (aproximadamente 80 viviendas). Dado que cada conglomerados está localizado en una área específica, la variabilidad **entre** conglomerados es mucho mayor para las áreas urbanas que en las áreas rurales, pero la variabilidad **dentro** del conglomerado seleccionado es menor en lo urbano que en lo rural. Por lo tanto, un importante factor en controlar el error muestral es reducir el mayor componente de variabilidad que es la variación entre conglomerados, el cual es inversamente proporcional al número de conglomerados (especialmente en el área urbana).

6. Influencia del nuevo marco muestral sobre los resultados de la encuesta

A partir del 2001 las ENAHO implementadas se han basado en muestras seleccionadas a partir del nuevo marco muestral de áreas cuya información básica proviene del Precenso de 1999. Este nuevo marco contiene aquellas áreas que han aparecido después del censo de población y vivienda de 1993. Estas áreas (pueblos jóvenes, asentamientos humanos, etc.) están localizadas básicamente en la periferia de las grandes ciudades o grandes aglomerados urbanos y constituyen un nuevo estrato el cual está altamente correlacionado con la pobreza extrema.

Es importante entonces, que en el nuevo marco muestral estén identificadas estas áreas de nuevo crecimiento a fin de que permita la comparabilidad de los resultados de las nuevas encuestas respecto a las encuestas anteriores cuyo diseño muestral se basaba en el marco anterior (Censo de 1993).

El INEI ha preparado una Tabla de Equivalencias para el área urbana, que permite comparar las manzanas identificadas en el precenso de 1999 versus las manzanas existentes en el Censo de 1993. Como resultado de esta comparación, puede conocerse la magnitud de los conglomerados del nuevo marco a nivel departamental, clasificados en manzanas coincidentes (que existen en el precenso del 99 y en el censo del 93), y las manzanas nuevas que representarían el incremento de la población urbana.

Al trabajar con la muestra nacional de la ENAHO del cuarto trimestre del 2001, se observa que el 77% de conglomerados urbanos son antiguos (ya existían en el censo de 1993), mientras que el 23% de los conglomerados urbanos son nuevos, los cuales representarían el incremento del área urbana en el periodo

de análisis. Tumbes es el departamento con mayor número de conglomerados nuevos, Moquegua ocupa el quinto lugar en el rango descendente de departamentos con un 39% de conglomerado nuevos.

El trabajo realizado para generar la tabla de equivalencias es de primera importancia porque permite que la selección de cualquier muestra que se seleccione con el nuevo marco, se puede replicar en la estructura del antiguo marco.

De este modo, una primera ronda de encuesta con el nuevo marco puede desagregarse en dos submuestras

a) la primera que contiene las USMs antiguas

b) la segunda submuestra que contiene las nuevas USMS,

Tal como se puede apreciar, la tabla es una herramienta que permite anular el efecto de diseño del nuevo marco muestral.

7. Muestras de panel de hogares y viviendas

Se conoce como Muestra Panel a aquella muestra en la que se miden los mismos elementos en dos o más periodos. Las encuestas por panel permiten hacer estudios de cambios individuales y de la dinámica de las causas y las relaciones.

En la Encuesta Nacional de Hogares (ENAHO), se ha venido utilizando paneles fijos con una rotación parcial de la muestra en el tiempo. Los paneles fijos son aquellos en los cuales los individuos que lo componen no cambian.

El uso exclusivo de un panel fijo produce estimaciones muy eficientes de cambios periódicos. En tanto que los paneles que rotan son utilizados para reducir el agotamiento de los respondientes y provee estimaciones eficientes de totales del periodo.

No ha habido un programa sostenido de rotación muestral y se han desarrollado varias experiencias. En la encuesta de 1996 se hizo una rotación del 25% de las viviendas de un trimestre a otro y solo duró un año. En 1997 se puso en práctica una muestra trimestral panel de visita anual, pero debido al pequeño tamaño de la muestra para una encuesta multipropósito, no hubo rotación hasta el año 2000. Esto debido a que en una evaluación en 1996 se encontró que el rechazo se presentaba en mayor medida a la segunda visita y luego esta se reducía en las posteriores. Además interesaba tener la mayor cantidad de viviendas en observación para poder detectar los determinantes de los cambios en las condiciones de vida. Este panel ha sido de viviendas y no de hogares, debido a la imposibilidad de detectarlas cuando los hogares cambian de residencia. Otro aspecto importante de remarcar es que en las visitas de seguimiento se entrevistaron a todas las viviendas seleccionadas en la primera ocasión, incluyendo aquellas que habían sido omisas, asegurando así que estos nuevos hogares se incorporan en la muestra, para su seguimiento posterior.

- **Encuesta Nacional de Hogares 1996**

El propósito principal de la ENAHO 1996 fue, mediante una muestra de tipo panel, medir los cambios en el tiempo del comportamiento de las características de las personas u hogares investigadas.

La muestra de 4 mil USM seleccionada para el primer trimestre de la ENAHO de 1995, fue utilizada en el primer trimestre de la ENAHO 1996. Es decir, se contó con un panel muestral de USM de año en año.

En la ENAHO DE 1996, en cada USM muestral se seleccionó una muestra de viviendas totalmente independiente a la seleccionada en 1995. La muestra nacional, de aproximadamente 20 mil viviendas, debía ser investigada durante los cuatro trimestres del año con una rotación parcial de un cuarto de la muestra de USM por trimestre, manteniendo un panel del 75 por ciento de hogares entre dos trimestres consecutivos.

Una desventaja de la encuesta por panel fue la reducción importante de la muestra debido a que los informantes se cansaron o perdieron interés, rehusándose a ser entrevistados.

DISTRIBUCIÓN Y RENDIMIENTO DE LA MUESTRA PANEL DE LA ENAHO 1996

VIVIENDAS	I TRIMESTRE	II TRIMESTRE	III TRIMESTRE	IV TRIMESTRE
Esperadas	20 000	20 000	20 000	20 000
Encuestadas	19 415	18 513	17 581	16 843
Rendimiento (%)	97,1	92,6	87,9	84,2

- **Encuesta Nacional de Hogares 1997**

En los primeros tres trimestres de 1997, la ENAHO estuvo orientada básicamente a la medición del empleo urbano. En el primero, segundo y tercer trimestre de la encuesta, la muestra para cada uno de ellos fue de 14 mil viviendas urbanas.

En cada trimestre se implementó una muestra panel de viviendas equivalente al 25 por ciento de la muestra global en cada uno de los trimestres. Los paneles de viviendas estaban referidos a las viviendas encuestadas en los trimestres respectivos de la ENAHO de 1996.

En el cuarto trimestre, se implementó una muestra urbano-rural, para efectuar la medición de los niveles de vida y pobreza de la población. En este trimestre no se utilizó muestra panel.

DISTRIBUCION DE LA MUESTRA DE LA ENAHO 1997

VIVIENDAS	I TRIMESTRE	II TRIMESTRE	III TRIMESTRE	IV TRIMESTRE
Seleccionadas: .Total . Panel	14 358 3 590	13 000 3 250	14 000 3 500	7 200 -
Rendimiento (%): . Total . Panel	90,5 81,1	91,9 82,2	87,5 84,0	89,8 -

- **Encuesta Nacional de Hogares 1998**

En el cuarto trimestre, se implementó una muestra urbano-rural de 7 mil 900 viviendas, para efectuar la medición de los niveles de vida y pobreza de la población.

En este trimestre, se implementó una muestra panel de viviendas con respecto al cuarto trimestre de 1997. Esta muestra panel fue de 5 mil 400 viviendas equivalente al 75 por ciento de la muestra total del trimestre.

DISTRIBUCION DE LA MUESTRA DE LA ENAHO 1998

VIVIENDAS	I TRIMESTRE	II TRIMESTRE	III TRIMESTRE	IV TRIMESTRE
Seleccionadas . Total . Panel	2 600 -	7 200 -	9 000 -	7 900 5 900
Rendimiento: . Total . Panel	92,6 -	93,6 -	92,9 -	86,3 84,7

- **Encuesta Nacional de Hogares 1999**

En el segundo trimestre de 1999, se implementó una muestra urbano-rural de 7 mil 600 viviendas a fin de obtener indicadores que permitan conocer la evolución de la pobreza, del bienestar y las condiciones de vida de los hogares.

En este trimestre, se implementó una muestra panel de viviendas con respecto al segundo trimestre de 1998. Esta muestra panel fue de 5 mil 700 viviendas equivalente al 75 por ciento de la muestra total del trimestre.

En el cuarto trimestre de 1999, se implementó una muestra urbano-rural de 4 mil viviendas a fin de obtener indicadores sobre la evolución de los niveles de vida y pobreza de la población.

En este trimestre, se utilizó la totalidad de la muestra de viviendas de la aplicada en el cuarto trimestre de 1998. Esta muestra panel de 4 mil viviendas era equivalente al 100 por ciento de la muestra total del trimestre.

DISTRIBUCION DE LA MUESTRA DE LA ENAHO 1999

VIVIENDAS	I TRIMESTRE	II TRIMESTRE	III TRIMESTRE	IV TRIMESTRE
Seleccionadas:				
. Total	7 200	7 600	9 000	4 000
. Panel	-	5 700	-	4 000
Rendimiento:				
. Total	92,8	86,7	90,8	85,8
. Panel	-	84,9	-	85,8

- **Encuesta Nacional de Hogares 2000**

En el segundo trimestre del 2 000, se implementó una muestra urbano-rural de 2 mil 720 viviendas a fin de obtener indicadores que permitan conocer la evolución de la pobreza, del bienestar y las condiciones de vida de los hogares.

En este trimestre, la totalidad de la muestra implementada fue un panel de viviendas con respecto al segundo trimestre de 1999. Esta muestra panel fue un equivalente al 35 por ciento de la muestra encuestada en el segundo trimestre de 1999.

En el cuarto trimestre del 2 000, se implementó una muestra urbano-rural de 4 mil 363 viviendas a fin de obtener indicadores sobre la evolución de los niveles de vida y pobreza de la población.

En este trimestre, se implementó una muestra panel de viviendas con respecto al cuarto trimestre de 1999. Esta muestra panel de 1 mil 800 viviendas equivale al 40 por ciento de la muestra total del trimestre.

DISTRIBUCION DE LA MUESTRA DE LA ENAHO 2000

VIVIENDAS	I TRIMESTRE	II TRIMESTRE	III TRIMESTRE	IV TRIMESTRE
Seleccionadas:				
. Total	1 300	2 720	12 000	4 363
. Panel	-	2 720	-	1 800
Rendimiento:				
. Total	86,1	82,6	93,2	83,9
. Panel	-	82,6	-	78,5

8. Uso de proyecciones de población para generar estimaciones

La metodología de estimación para procesar los datos de la ENAHO, involucra el uso de un peso o factor de expansión para cada registro que será multiplicado por todos los datos que conforman el registro correspondiente. El factor final para cada registro tiene dos componentes: el factor básico de muestreo y los factores de ajuste por la no entrevista.

El factor básico de expansión para cada hogar muestral es determinado por el diseño de la muestra bajo la hipótesis de que todas las viviendas seleccionadas en el conglomerado que salen a campo, retornarán todas ellas con información. Equivale al inverso de su probabilidad final de selección, el mismo que es el producto de las probabilidades de selección en cada etapa. El factor básico

El factor de expansión consiste en la asignación de un factor o coeficiente a cada uno de los hogares, de tal forma que aplicado a la muestra la información obtenida nos represente el universo (expandida), es así como se construye el factor "Primario" de expansión en función a:

Luego de finalizada la operación de campo se procede a realizar el ajuste del factor básico de expansión en función a 2 criterios:

- El ajuste por la Tasa de No Respuesta, que consiste en evaluar en cada conglomerado el total de viviendas con información (con resultado completo e incompleto), luego de conocida esta información se procede a reajustar el factor de expansión.
- El ajuste por la población, se considera como premisa la proyección de población para el año que se está trabajando y se reajusta el factor de expansión hasta obtener valores que estén alrededor de este dato.

D. Aspectos Operativos para la Construcción del Marco Muestreo

1. Creación de áreas de listado (diferencias entre áreas urbanas y rurales)

Censalmente, se conoce como centro poblado a todo lugar del territorio nacional, identificado mediante un nombre y **habitado con ánimo de permanencia al cual los habitantes se encuentran vinculados** por intereses comunes de carácter económico, social, cultural e histórico. La categoría de Centro Poblado es la denominación que la ley le asigna o la clase o rango con la que le denomina o conoce.

En el Censo de 1993 los centros poblados urbanos eran las urbanizaciones, barrios, asentamientos humanos, pueblos jóvenes, etc. En 1999 la definición se modifica. Las agregaciones de los núcleos urbanos formaron los Centros Poblados, varios núcleos urbanos por agregación forman un centro poblado.

Para efectos censales, el centro poblado urbano es aquel que tiene como mínimo 100 viviendas agrupadas contiguamente y todos los centros poblados que son capitales de distritos, aún cuando no reúnan la condición indicada. Por otro lado, el centro poblado rural es aquel que no tiene más de 100 viviendas contiguamente ni es capital de distrito, o que teniendo más de 100 viviendas, estas se encuentran dispersas o diseminadas sin formar bloques o núcleos

Para efectos de la Encuesta Nacional de Hogares, el centro poblado urbano es aquel que tiene como mínimo 2 mil habitantes, con servicios de agua, desagüe, alumbrado público y cuya actividad económica preponderante no pertenece al sector primario. Mientras que el centro poblado rural es aquel que tiene menos de 2 mil habitantes, sin servicios básicos completos o solamente tienen algún servicio que atienden parcialmente a sus habitantes; la actividad económica preponderante pertenece al sector primario.

2. Procedimientos operativos para la elaboración de listados de viviendas

Desde la ejecución del Precenso de 1999 a la fecha, el marco muestral de la encuesta ha sufrido un desgaste por "envejecimiento", lo que puede ocasionar, una vez implementada el operativo de recopilación de datos por los encuestadores, se tenga altas tasas de no-respuesta debido a la falta de actualización de los directorios.

En consecuencia, ha sido necesario efectuar una actualización del directorio de viviendas en cada conglomerado seleccionado para la muestra. Esta tarea se desarrolla de la manera siguiente:

- Personal especialmente entrenado, con ayuda del listado de edificios y viviendas, recorre íntegramente todo el perímetro de cada manzana que compone el conglomerado seleccionado.
- Durante este recorrido toman nota de las modificaciones que se hayan producido tales como cambio en la forma de la manzanas, apertura o cierre de calles, nombre en calles que no lo tenían, cambio en dichos nombres, cambios en la numeración de las viviendas, cambios en el tipo de vivienda, detección de errores en el listado original, etc.
- Incorpora las viviendas nuevas y elimina con una línea los errores o las viviendas demolidas.
- Después de la actualización, mediante trabajo de gabinete, se realiza una evaluación del tamaño de cada conglomerado. Esto, para evitar que en el marco muestral se presenten conglomerados con un número atípico de viviendas.
- El tamaño promedio del conglomerado es de 80 viviendas. Así, al localizarse un conglomerado con más de 120 viviendas, se procede a su división, y si el número de viviendas es menor a 40, se procede a fusionarlo con otro adyacente o contiguo. Esto da origen a que en el marco muestral aparezcan, adicionalmente, los "conglomerados divididos" y "conglomerados fusionados" como casos especiales.
- En el caso de los conglomerados divididos, se le agrega a la identificación original un número consecutivo que empieza en 1 y continúa hasta terminar en un número igual al número de compactos en que se dividió el conglomerado.

En el caso de los conglomerados fusionados, se agrega a la identificación original, el numero del conglomerado con el cual se fusionó. Al fusionarse un conglomerado, la identificación del primero permanece en el marco y la otra desaparece.

3. Actualización permanente del marco de muestreo y de la muestra

El INEI considera que la actualización sistemática del marco muestral de áreas debe ser una política permanente de tal manera que las muestras obtenidas de los marcos maestros puedan ser complementadas con selecciones de viviendas en las áreas nuevas, para asegurar así la real representación de la población. Se ha tenido experiencias negativas de que la no actualización del marco conduce a la obtención de indicadores sesgados, al producirse un desvalance en los pesos urbanos-rural y los asentamientos por niveles de pobreza, pues el crecimiento urbano en la periferia de las ciudades de los últimos 10 años ha estado constituido mayormente por poblaciones muy pobres. El marco muestral de áreas actual de la ENAHO será actualizado a partir de la Base de Datos y la Cartografía digitalizada de los Censos Nacionales: X de Población y V de Vivienda a levantarse con el 2004.

El marco de muestreo utilizado por la ENAHO desde 1995 hasta el 2000, fue actualizado en 1996 en las grandes ciudades del país. Esta tarea tuvo como objetivo principal incorporar en el marco las áreas de nuevo crecimiento. El trabajo de actualización se efectuó en Lima Metropolitana, ciudad capital del país, Ciudades Capitales Departamentales y otras Grandes Ciudades del país.

4. Tratamiento de viviendas, adicionales, omitidas y desaparecidas

En la Encuesta Nacional de Hogares en la última etapa de muestreo se selecciona un número fijo de viviendas ya sea mediante una selección sistemática simple aleatoria o bien seleccionando segmentos compactos de viviendas. En ambos casos, los listados de viviendas seleccionadas sirven para que el encuestador identifique las viviendas omitidas o adicionales al registro. Estas viviendas no son entrevistadas pero sí enumeradas. Si son un número importante, se procede a la actualización inmediata.

5. Tratamiento de estratos especiales de viviendas

En la Encuesta Nacional de Hogares no son investigados locales de alojamiento especial como son los hospitales, hoteles, hostales , comisarías , puestos policiales, cuarteles, internados, claustros religiosos, hospitales, etc.

Así mismo, en el área urbana existen grupos de viviendas especiales como son las villas militares, grupos de viviendas en asentamientos mineros. Estos tipos de viviendas están conformadas en la mayoría de los casos por hogares unipersonales u hogares en tránsito por motivo de trabajo. Estos hogares son investigados siempre y cuando cumplan requisitos de residencia habitual. En el caso de las villas militares la colaboración a las encuestas de hogares es limitada, demanda tiempo completar las entrevistas y el rechazo es bastante frecuente.

MARCO MUESTRAL DE PROPÓSITOS MÚLTIPLES Y MUESTRA MAESTRA

BANCO CENTRAL
REPÚBLICA DOMINICANA

ÍNDICE

Página

Introducción 293
I.- Marco Muestral de Propósitos Múltiples 293
1.1 Cobertura 293
1.2 Estructura censal 293
1.3 Formación de estratos 294
1.4 Marco muestral de edificaciones nuevas 295
II.- Muestra Maestra 296
2.1 Diseño muestral de la muestra maestra 296
2.2 Formación de las réplicas en la muestra maestra 297

Introducción

Los dos aspectos básicos que se toman en consideración en el planeamiento del diseño estadístico de una encuesta de hogares, lo constituyen la construcción del marco muestral, que abarca un listado de todas las unidades de la población de interés que participarán en las distintas fases de selección de la muestra, el procedimiento aplicado para la selección de las unidades de observación de la muestra y la definición de los estimadores.

I. Marco Muestral de Propósitos Múltiples

La necesidad de construir el Marco Muestral de Propósitos Múltiples (MMPM) surge, principalmente, del agotamiento del marco muestral existente basado en el Censo de Población y Vivienda de 1981 y del costo de su actualización. En efecto, construir un nuevo marco, basado en la cartografía más reciente proveniente del Censo Nacional de Población y Viviendas de 1993, resultaba menos costoso que actualizar el marco existente.

Por esta razón y con el objetivo de elaborar diseños para las Encuestas de Fuerza de Trabajo (EFT) y la Tercera Encuesta Nacional de Gastos e Ingresos de los Hogares (ENGIH), el Banco Central inicia en el año 1994 los trabajos relativos al proyecto de construcción del MMPM, el cual contó con la asistencia técnica y financiera de la Organización Internacional de Trabajo (OIT).

1.1 Cobertura

El MMPM abarca el total de viviendas no colectivas ocupadas, quedando excluidas las viviendas colectivas tales como, hospitales, hoteles, moteles, apartahoteles, cárceles, conventos, escuelas, asilos, auspicios, cuarteles, entre otros; así como, los hogares individuales ubicados en viviendas colectivas.

La cobertura geográfica del MMPM tiene alcance nacional, es decir, que las viviendas no colectivas ocupadas de las zonas urbanas y rurales de todos los municipios son susceptibles de ser seleccionadas con apropiados niveles de probabilidad.

1.2 Estructura censal

La construcción del MMPM se hizo en base a la división político administrativa del Censo Nacional de Población y Vivienda de 1993, realizado por la Oficina Nacional de Estadísticas, utilizando un listado de todas las unidades geográficas denominado Hojas de Segmentación (HS), el cual contenía la identificación de las siguientes unidades censales:

- Provincia
- Municipio
- Barrio o Sección
- Polígono
- Area de Supervisión
- Segmento

En adición a los códigos que identifican cada una de estas unidades, las HS contenían además, la cantidad de hogares y viviendas en la unidad censal más pequeña, es decir, el segmento.

La información contenida en las HS fue revisada por muestra, encontrándose discrepancias no admisibles entre el número de hogares y de viviendas existentes en la unidad segmento y duplicaciones en los códigos de identificación de los segmentos, problemas éstos que son usuales en la mayoría de los censos. Por ello, se realizó una cuidadosa depuración de todas las unidades contenidas en las hojas de segmentación.

Asimismo, se determinaron hojas de segmentación faltantes y la existencia de segmentos sin viviendas. En este último caso, con la ayuda de los técnicos de la ONE, se asignó una cantidad de vivienda a cada uno de los segmentos para garantizar su inclusión probabilística en la selección de la muestra. El procedimiento se basó en asignar un número de vivienda por segmento, en alguno casos, de acuerdo al promedio de viviendas de otros segmentos de la misma área de supervisión y en otros, en función de la carga de trabajo censal de un empadronador.

Tomando en consideración todo lo anterior, en la selección probabilística de las unidades primarias de muestreo (UPM) se asumió utilizar los hogares, como medida del tamaño para la selección.

1.3 Formación de estratos

Las informaciones disponibles de las HS se refieren a la división político administrativa del país, por lo que no permitieron formar estratos socioeconómicos diferenciados con base a características de las viviendas, los hogares que la habitan y sus miembros. Para fines de la ENGIH las provincias con sus respectivos municipios se agruparon en regiones de acuerdo a su importancia poblacional y/o cercanía geográfica, las cuales se les denominó estratos.

Adicionalmente, a la zona urbana del Distrito Nacional y el municipio de Santiago, donde se encuentran las mayores concentraciones poblacionales del país, se les agregaron secciones rurales, cuya población tiene características urbanas; por esta razón, se les ha denominado, Santo Domingo y Santiago Metropolitano. Cabe señalar por tanto, que las zonas urbanas y rurales de las regiones o estratos construidos, no son comparables con otros diseños de investigación que estén configurados estrictamente con los lineamientos de la división política territorial del país. Por tanto, este marco muestral corresponde al denominado de áreas, porque las unidades que lo configuran son unidades geográficas.
En la tabla siguiente se describen las regiones o estratos geográficos y las provincias que la conforman.

Distribución Geográfica de las Provincias Según Regiones

REGIONES	PROVINCIAS
SANTO DOMINGO METROPOLITANO	Incluye todos los barrios urbanos de Santo Domingo y • Sección Cancino • Sección Haina de Engombe • Paraje Guaricano de la Sección Higueral • Parajes Los Alcarrizos y Pantoja de la Sección Los Alcarrizos • Parajes San Miguel, Manoguayabo y Bayona de la Sección de Manoguayabo • Parajes Mendoza, San Miguel, Hainamosa y La Cana de la Sección Mendoza Paraje Villa Mella de la Sección Villa
SANTO DOMINGO RURAL	Secciones de: Boca Chica, El Coco de Pedro Brand, Guerra, Hato Viejo, Higuero, La Bomba, La Victoria, Pedregal Los Alcarrizos, Manoguayabo. Mendoza y Villa Mella (con excepción de las secciones y parajes que pasan a la parte urbana.
SANTIAGO METROPOLITANO	Incluye todos los barrios urbanos del municipio de Santiago, y • Paraje Canabacoa de la Sección Canabacoa • Parajes Ingenio Arriba, Campeches y Rafey de la Sección El Ingenio Abajo • Paraje Arroyo Hondo de la Sección El Papayo • Paraje Gurabo Arriba de la Sección Gurabo Abajo Paraje Jacagua de la Sección Jacagua
SANTIAGO RURAL	Incluye el resto de las secciones rurales del municipio Santiago excepto las clasificadas como urbanas.
CIBAO URBANO	Incluye las provincias de: Montecristi, Dajabón, Santiago Rodríguez, Valverde, Puerto Plata, La Vega, Monseñor Nouel, Espaillat, Salcedo, Duarte, Sánchez Ramírez, María Trinidad Sánchez, Samaná y los municipios urbanos de la provincia de Santiago.
CIBAO RURAL	Incluye las zonas rurales de las mismas provincias del renglón anterior.
SUROESTE URBANO	Incluye las provincias de: San Cristóbal, Peravia, Azua, San Juan de la Maguana, Barahona, Bahoruco, Independencia, Elías Piña y Pedernales.
SUROESTE RURAL	Incluye las zonas rurales de las mismas provincias del renglón anterior.
SURESTE URBANO	Incluye las provincias de: Monte Plata, San Pedro de Macorís, Hato Mayor, El Seybo, La Romana y La Altagracia.
SURESTE RURAL	Incluye las zonas rurales de las mismas provincias del renglón anterior.

La distribución de los segmentos y las viviendas por regiones correspondientes al MMPM se muestran en la siguiente tabla:

Distribución de los Segmentos Censales y las Viviendas del MMPM, Según Regiones o Estratos

Regiones	Total Segmentos	Total Viviendas
Santo Domingo Urbano	9,568	381,699
Santo Domingo Rural	823	49,778
Santiago Urbano	1,976	80,900
Santiago Rural	735	23,582
Cibao Urbano	4,440	169,569
Cibao Rural	10,082	241,318
Suroeste Urbano	3,329	123,970
Suroeste Rural	5,098	142,186
Sureste Urbano	2,605	97,878
Sureste Rural	3,339	70,727
TOTAL	**41,995**	**1,381,607**

1.4 Marco muestral de edificaciones nuevas

La existencia de urbanizaciones y/o edificaciones, surgidas con posterioridad a la finalización del Censo de Población y Vivienda de 1993, conllevó a la construcción del Marco Muestral de Edificaciones Nuevas (EN). Este nuevo marco se construyó sólo para Santo Domingo y Santiago metropolitanos, ciudades que mantienen un acelerado ritmo de construcción. Los criterios considerados para la inclusión de estas edificaciones en el mismo, fueron:

1. Que fueran edificaciones tipo apartamento y/o urbanizaciones.
2. Que la urbanización o edificación nueva no formara parte de las UPM seleccionadas para evitar duplicaciones.
3. Que con anterioridad a la construcción de la edificación y/o la urbanización existiera un solar baldío.
4. Que la construcción de la urbanización o edificación nueva se hubiera realizado entre los años 1994 y 1996.
5. En ningún caso, se incluyeron en la EN viviendas del tipo casa individual.

Para la construcción del marco muestral de edificaciones nuevas Santo Domingo metropolitano se creó una base de datos partiendo de los listados de las urbanizaciones y edificios donde se habían instalado contadores de consumo de energía eléctrica en los años comprendidos entre 1994 y 1996, suministrados por las Oficinas Comerciales de la Corporación Dominicana de Electricidad. Las urbanizaciones o edificaciones incluidas en estos listados fueron depuradas mediante visitas a las mismas, lo que permitió comprobar el cumplimiento de los criterios de inclusión en dicho marco.

Estos listados fueron enriquecidos con las observaciones del personal de oficina que laboró en la construcción del MMPM, los que informaron sobre edificaciones no clasificadas en el listado anterior o en las hojas de segmentación del censo. En el caso de Santiago metropolitano, este mismo personal aportó las informaciones para la construcción del marco de edificaciones nuevas. La cantidad de viviendas que integran este marco se presenta a continuación:

Marco Muestral de Edificaciones Nuevas Distribución de las Viviendas

Estratos	Total Viviendas
Santo Domingo	6,970
Santiago	2,315
TOTAL	**9,285**

II. Muestra Maestra

Con base al marco muestral descrito, se establecieron las condiciones para construir la Muestra Maestra, definida como aquella a partir de la cual pueden seleccionarse muestras independientes o réplicas, con la finalidad de ser usada para diferentes propósitos, a saber la Encuesta de Fuerza de Trabajo (EFT) y la Encuesta Nacional de Gastos e Ingresos de los Hogares (ENGIH).

2.1 Diseño muestral de la muestra maestra

El diseño muestral de la muestra maestra utilizado fue del tipo bietápico, estratificado (regiones geográficas) de conglomerados con selección proporcional al tamaño en la primera etapa y selección sistemática-aleatoria en la segunda etapa. Este procedimiento muestral, dispone de cinco réplicas seleccionadas aleatoriamente y de forma independiente en cada una de las regiones geográficas o dominios de interés.

La primera etapa del diseño consistió en la selección de las Unidades Primarias de Muestreo (UPM), constituidas por aproximadamente 60 viviendas, siguiendo el siguiente procedimiento:

- Se ordenaron los municipios en forma serpentina para cada una de los estratos y los barrios en el caso de Santo Domingo y Santiago Urbano, y para cada uno de ellos se obtuvo un acumulado de hogares.
- Se procedió a la selección de las UPM, con base al criterio de probabilidad proporcional al tamaño (ppt), definiéndose por tamaño el número de viviendas ocupadas existente en cada uno de los segmentos censales.
- Se procedió a la actualización exhaustiva de cada una de las UPM, lo cual consistió en identificar la ubicación geográfica del segmento seleccionado que provenía de la Oficina Nacional de Estadística (ONE) y verificar el número de viviendas existentes en el mismo.

En la segunda etapa, se realizó la selección sistemática- con arranque aleatorio- de las viviendas partiendo de las UPM seleccionadas y actualizadas. La experiencia operativa de encuestas de Fuerza de Trabajo anteriores a la construcción de esta Muestra Maestra, indicaba que un encuestador podía realizar aproximadamente 10 entrevistas en un día, por tanto se seleccionó en cada segmento muestral 10 viviendas, es decir que se utilizó como criterio la carga de trabajo de un encuestador. Conviene señalar que este número de viviendas puede variar en función del tipo de investigación.

Los tamaños muestrales en las diferentes regiones, se muestran a continuación:

Marco Muestral de Propósitos Múltiples
Distribución de las UPM por Estratos

Regiones	Cantidad UPM	Viviendas por UPM
Santo Domingo Metropolitano	180	1,800
Santo Domingo Rural	21	210
Santiago Metropolitano	51	510
Santiago Rural	9	90
Cibao Urbano	60	600
Cibao Rural	69	690
Suroeste Urbano	42	420
Suroeste Rural	48	480
Sureste Urbano	36	360
Sureste Rural	36	360
TOTAL	**552**	**5,520**

2.2 Formación de las réplicas en la muestra maestra.

La réplica es una muestra aleatoria representativa de la población del estrato del cual proviene. En un diseño muestral pueden utilizarse en la rotación de las UPM, en la evaluación de la calidad de los encuestadores y supervisores y en la flexibilidad hacia otros tipos de diseños muestrales, entre otros.

La Muestra Maestra dispone de 5 réplicas para cada uno de los estratos. El procedimiento utilizado para determinar los arranques aleatorios de las 5 réplicas es el siguiente:

- Se estima la cantidad de UPM por estrato en una muestra de cinco réplicas, tomando en consideración la experiencia de otros diseños en países de la región.

Distribución de las UPM en la Muestra Maestra

Regiones	Cantidad UPM en la Muestra Maestra	Replicas de la Muestra Maestra
Santo Domingo Urbano	300	5 de 60
Santo Domingo Rural	35	5 de 7
Santiago Urbano	85	5 de 17
Santiago Rural	15	5 de 3
Cibao Urbano	100	5 de 20
Cibao Rural	115	5 de 23
Suroeste Urbano	70	5 de 14
Suroeste Rural	80	5 de 16
Sureste Urbano	60	5 de 12
Sureste Rural	60	5 de 12
TOTAL	**920**	

- Se determina el arranque aleatorio de selección de las UPM, para lo cual se forman cinco categorías o clases con un tamaño dado por el cociente del total de viviendas en el MMPM entre el número de UPM en la muestra maestra de 5 réplicas. Por ejemplo, para el caso de Santo Domingo Urbano el tamaño de cada clase se determina de esta manera: los 381,699 hogares que este estrato tiene en el MMPM se dividen entre las 300 UPM existente en la muestra maestra de cinco réplicas para el mismo estrato, lo que arroja un resultado de 1,272, es decir, esta sería la amplitud de las cinco clases o categorías.

- El arranque aleatorio, que es el punto de partida para la selección de las UPM se escogió de un número aleatorio (k) entre 1 y 1,272, el cual fue de 683. Los demás arranques aleatorios para las demás réplicas se obtienen sumando el valor k al límite superior de la clase anterior. El siguiente cuadro ilustra los dos pasos explicados.

Arranques Aleatorios de la Muestra Maestra

Por Intervalos de Clases y Réplicas Clases	Réplicas	Arranques Aleatorios
0001 – 1,272	Primera	683
1,273 – 2,545	Segunda	1,272 + 683 = 1,955
2,546 – 3,818	Tercera	2,545 + 683 = 3,228
3,819 – 5,091	Cuarta	3,818 + 683 = 4,501
5,092 – 6,364	Quinta	5,091 + 683 = 5,774

Esta forma de determinación del arranque aleatorio de las cinco réplicas se utilizó con la finalidad de garantizar la dispersión entre las unidades seleccionadas entre las réplicas y su condición de independencia.

En la selección de las UPM en cada réplica se parte del número del arranque aleatorio en su correspondiente réplica y de un intervalo de selección calculado por el cociente del total de hogares en el MMPM y el número de UPM a seleccionar en cada réplica. Para el mismo ejemplo de Santo Domingo urbano, los 381,699 hogares en el MMPM se dividen entre las 60 UPM a seleccionar en dicho estrato, lo que arroja un resultado de 6,362 hogares; esto indica que de cada 6,362 hogares, se seleccionan las UPM de cada réplica. [1]

En resumen, para las cinco réplicas de la muestra maestra la selección de las UPM sería la siguiente:

1ra. Réplica:	683;	683 + 6,362;	7,045 + 6,362; ...
2da.Réplica:	1,955;	1,955 + 6,362;	8,317 + 6,362; ...
3ra. Réplica:	3,228;	3,228 + 6,362;	9,590 + 6,362; ...
4ta. Réplica:	4,501;	4,501 + 6,362;	10,863 + 6,362; ...
5ta. Réplica:	5,774;	5,774 + 6,362;	12,136 + 6,362; ...

Debido a la dificultad y al costo de actualizar las UPM seleccionadas para todas las réplicas de la Muestra Maestra, se realizó un trabajo de campo de actualización de las UPM seleccionadas en las tres réplicas reservadas para las Encuesta de Fuerza de Trabajo. Esta revisión arrojó que la mayoría de los segmentos seleccionados contenían límites imaginarios, y en otros casos, éstos habían desaparecidos, lo que obligó a delimitar las UPM partiendo de límites naturales y estables en el tiempo. Asimismo, se observaron discrepancias entre la cantidad de viviendas registradas en las hojas de segmentación y en la cantidad de viviendas de los croquis de los segmentos.

[1] La diferencia presentada entre el cociente 6,362 y el límite superior del intervalo de la última clase de 6,364, es debido a los efectos de redondeo.

DISEÑO DE LA MUESTRA DE LA ENCUESTA CONTINUA DE HOGARES

INSTITUTO NACIONAL DE ESTADÍSTICA (INE)
URUGUAY

ÍNDICE

Página

Objetivos Generales de la Encuesta Continua de Hogares (ECH) 301
1.- Necesidades del Rediseño de la Muestra 301
2. - Breve Historia de la ECH 303
3. - Niveles Geográficos de Estimación 305
4. - Localidades del Interior Urbano y de la Periferia 306
5. - Períodos de Desagregación 309
6. - La Estratificación Vigente 309
7. - Tamaño de la Muestra Anual 311
Anexo 1 División Geográfica y Estadística del País 317

Se ha considerado adecuado remitir el presente documento ya que constituye el principal uso que se le ha dado al marco confeccionado a partir del Censo General de Población, Viviendas y Hogares levantado en el año 1996 en el Uruguay.

Se deja expresa constancia que el presente Documento fue preparado por M. Mónica Beltrami, al que se le practicaron puequeñas modificaciones para operativizarlo para el 9° Taller Regional del Programa MECOVI.

Objetivos Generales de la Encuesta Continua de Hogares

Como ha sido habitual, la ECH tiene por objetivo primario estimar la *tasa de actividad económica*, de *desocupación* y el *ingreso total medio de los hogares* en áreas seleccionadas del territorio nacional.

Como objetivo secundario, la ECH permite la estimación de las relaciones o diferencias de las tasas de actividad y desocupación según otras variables de control como sexo, parentesco, tipo de desocupación, rama de actividad, ocupación, categoría de la ocupación y nivel de instrucción, entre otras.

El ingreso del hogar, por su parte, puede estimarse clasificándose por algunas de esas mismas variables de control, o expresarse por decil/quintil/cuartil, *per cápita*, o por sus componentes (ocupación principal de las personas, fuente de ingreso, y otras). En las entrevistas e informes de coyuntura se expresan en moneda corriente, pero también pueden y deben recalcularse en moneda real [sea Unidades Reajustables (UR), dólares norteamericanos, salarios mínimos o cualquier otro tipo de valor deflactado], con propósitos comparativos.

La misma encuesta sirve para estimar otras características de la población objetivo, vinculadas -o no- con las variables listadas antes. Por ejemplo, cuando se incluyen en el cuestionario temas adicionales pedidos por usuarios o relevados como estudio periódico de características estructurales de la población, las variables medidas en estos módulos pueden vincularse -o no- con las variables primarias de inclusión continua.

1. Necesidad del Rediseño de la Muestra

Al relevarse el VII Censo General de Población, III de Hogares y V de Viviendas (Censo 1996) el 22 de mayo de 1996, ya se había estimado necesario replantear la ECH en términos del nuevo marco muestral (el Marco 1996) y de un estudio más profundo y extenso de estratificación, subpoblaciones y niveles de estimación por área, período o variables de control.

Es así que la nueva muestra de la ECH 1998 incluye:

a) una nueva estratificación en Montevideo,

b) la extensión del relevamiento de Montevideo a su área periférica (una especie de "Gran Montevideo"),

c) la posibilidad de producir resultados a nivel de *departamento* (ver definición en el Apéndice 1) del resto urbano del país,

d) la definición del término "urbano" asignándolo a las *localidades o centros poblados* (definición en el Apéndice 1) del Interior que totalizaron 5.000 habitantes o más (al Censo 1996), y

e) una optimización de tamaños muestrales tanto para el área extendida de Montevideo como para el Interior urbano.

La nueva muestra constituye, en primer lugar, un ejercicio de puesta en uso del Marco 1996, y su evaluación a campo con relación al Marco 1985 y sus alternativos: el marco de listado de contadores de UTE o el listado de "carpetas catastrales" de la Intendencia Municipal de Montevideo (IMM) para este último departamento. En segundo lugar, consiste en la actualización de tamaños muestrales a los niveles geográficos (Montevideo,

Interior urbano, etc.) y la formulación de un diseño muestral más apropiado para la estimación de las variables e indicadores que se estudian en forma continua.

1. Etapas del trabajo

Como primer paso, se estableció un programa de trabajo para:

a) definir los niveles geográficos de estimación y sus desagregaciones: estratos socio-económicos y de tamaño.

b) conseguir un marco muestral a nivel de *zona censal* [*unidad de primera etapa* de selección, típicamente una manzana (definición en el Apéndice 1)], para los principales niveles geográficos de estimación;

c) determinar el/los parámetro/s cuya estimación es de principal importancia;

d) obtener las estimaciones en los niveles geográficos mayores existentes en la ECH (91-96) aplicando el presunto diseño vigente con todas sus especificaciones para diversos períodos usuales de estimación (mes, trimestre, semestre);

e) calcular el tamaño muestral para el menor período de desagregación en cada nivel geográfico, y calcular los coeficientes de variación esperados por agregación a los períodos de publicación vigentes;

f) seleccionar los menores niveles geográficos (no necesariamente de desagregación) dentro del llamado *Interior urbano*, esto es, las *localidades o centros poblados* en las que se seleccionarán las unidades de información;

g) distribuir y ajustar los tamaños muestrales a las desagregaciones geográficas definidas en a) y f);

h) calcular la cantidad de zonas (unidades de primera etapa de selección) en cada localidad y agrupamiento;

i) verificar que las zonas a ser seleccionadas en cada localidad tengan 3 o más viviendas (*unidades de segunda etapa* de selección). Si esto no ocurre, agrupar las zonas con menos de 3 viviendas a las zonas vecinas del mismo *segmento censal* (definición en el Apéndice 1) hasta que sí agrupen 3 o más viviendas;

j) proceder a la selección de zonas, con probabilidad proporcional a su tamaño.

2. Desarrollo del trabajo

Los niveles geográficos de desagregación se definieron en función de la experiencia recogida en el relevamiento de la ECH de los últimos 17 años, de las demandas de información por parte de los usuarios (en la medida de lo razonable en lo que tiene que ver con la confiabilidad de las estimaciones) y de la información *volumétrica* proveniente de los recuentos a nivel de "zona censal" del Censo.

El plan de trabajo descripto en A, se desarrolló en dos instancias: la primera entre aproximadamente el 15 de agosto y fines de octubre de 1996. La segunda entre noviembre de 1996 y octubre de 1997.

En la primera instancia se logró establecer el tamaño muestral óptimo y completar como ejercicio la selección de unidades de primera y de segunda etapa. Puesto que solamente se contaba con un marco proveniente de los *Recuentos Preliminares* del censo construido con las hojas de control de empadronamiento de los Jefes de Segmento (documento censal codificado *C9*), este ejercicio no se cristalizó en una corrida a campo de la encuesta.

En la segunda instancia, ya con el Marco 1996 "definitivo", se sortearon las localidades urbanas, se distribuyeron y ajustaron los tamaños muestrales, se reiteró la selección de unidades de primera y segunda

etapa y se las distribuyó a los períodos de referencia (semanas) de tal modo que todas las localidades y estratos estuviesen igualmente representados a lo largo del año.

2. Breve Historia de la ECH

La Encuesta de Hogares (ECH) comenzó a relevarse en 1968 en forma continuada. Hasta 1979 su alcance geográfico permanente fue Montevideo, pero aunque demandas de usuarios y recomendaciones del I Seminario Estadístico Nacional de 1970 lo reclamaron, y hubo intentos positivos de extensión a otros departamentos (por ejemplo, Canelones en 1976), estas iniciativas tuvieron poco éxito debido a escasez de recursos tanto humanos como presupuestales.

Por fin en 1980, gracias a un convenio con el Programa de las Naciones Unidas para el Desarrollo y el Fondo de las Naciones Unidas para Actividades de Población, se obtuvo apoyo y recursos para esa extensión a todo el país y asimismo se re-diseñó la muestra, incluyendo los cuestionarios que, con algunas pequeñas modificaciones estructurales, se vienen aplicando hasta la fecha.

También por esa época la "*revolución informática*" alcanzó a la entonces Dirección General de Estadística y Censos, con la adquisición del primer *mini-computador* que contó este Instituto, para el procesamiento de sus crecientes volúmenes de datos.

1. La cobertura geográfica

Las áreas relevadas desde 1980 hasta hoy fueron y son: Montevideo e Interior Urbano. [1]

El alcance de la definición de lo "urbano" y de la cobertura geográfica (aun en Montevideo), experimentaron cambios a lo largo del tiempo.

1. El Interior

Al principio, lo "urbano" del Interior cubrió los centros poblados de 5.000 o más habitantes al Censo de 1975. Por razones operativas esos centros poblados fueron representados por las 18 capitales departamentales y dos ciudades mayores del área periférica de Montevideo: Las Piedras y Pando. Nos referiremos a todas ellas como "*las 20 mayores*".

Entre 1982 y 1985 se extendió la cobertura urbana a los centros poblados de 900 a 5.000 habitantes al Censo 1975, y entre 1986 y 1990 a las correspondientes en ese estrato de tamaño al Censo 1985. En ese estrato se seleccionaron al azar las localidades en las que se tomaría la encuesta. La lista de estas localidades varió poco año a año.

En 1991 con un cambio en el cuestionario base, se revisó también la cobertura geográfica del Interior. Las 20 mayores alcanzarían a representar los centros poblados de 10.000 o más habitantes y una muestra de localidades menores cubriría el estrato de 900 a 10.000. [2] Al año siguiente, un replanteo de fondo dio peso proporcional a cada uno de tres estratos de tamaño: *mayor* (de 10.000 o más), *mediano* (de 5.000 a 10.000) y *pequeño* (900 a 5.000), este último a su vez sub-estratificado en 5 grupos según algunos indicadores socio-económicos y geográficos que se consideraron adecuados.

Entre 1993 y 1997 se estableció la cobertura urbana del Interior que se ha mantenido por más tiempo: las 20 mayores y 32 centros poblados de 900 a 10.000 habitantes. [3]

---

[1] En 1981 la ECH relevó por única vez el área rural del país. La muestra tuvo un volumen total de 2.400 hogares y la publicación de sus resultados se produjo en 1982. Como se explica más adelante fueron varios factores, entre ellos la falta de recursos económicos, que impidieron continuar el relevamiento de áreas rurales del país.

[2] Ver : **URUGUAY: Informe de Misión**, Edmundo Berumen, Consultor, CEPAL. Marzo 1992.

[3] Ver : Berumen, E. (informe citado).
Metodología de la Encuesta Continua de Hogares. DGEyC, 1991-1993.
Encuesta Nacional de Hogares. DGEyC, publicaciones anuales de 1981 a 1990.
URUGUAY. Marco muestral y métodos de Selección de la Muestra de Viviendas, Informe de Misión de Carlos Cavallini, Asesor, FNUAP. Diciembre 1978.

2. Montevideo

Variaciones no tan acusadas sufrió la cobertura geográfica de Montevideo. En general, la ECH cubrió todo el departamento, incluyendo los segmentos rurales, los cuales tienen características bastante más "urbanas" que sus homónimos del resto del país. Entre 1984 y 1989 se excluyeron estos segmentos rurales por diversas consideraciones operativas.

2. La estratificación

Desde 1980, Montevideo y el Interior urbano fueron estratificados en forma separada y por distintos criterios.

1. Estratificación del Interior urbano

Con la estratificación por proximidad geográfica (6 "regiones") que se estableció en el diseño de 1980 y cuya existencia fue documentada en la metodología de las encuestas de 1981-84 y 1986-90, nunca se produjeron estimaciones a nivel "región" pero sí a nivel "departamento" a expreso pedido de los usuarios. [4]

Como curiosidad, cabe citar que estas regiones geográficas fueron denominadas explícitamente "estratos" hasta 1986. Desde la publicación anual de 1987 las regiones pasaron a designarse "divisiones operativas".

La afijación de los tamaños muestrales fue, en general, *proporcional al tamaño de los estratos de centros poblados*, medido en cantidad de viviendas particulares al censo más reciente.

Las unidades de primera etapa de selección del Interior fueron siempre las *zonas censales* o "manzanas" o "blocks", seleccionados con probabilidad proporcional a su tamaño. Las unidades de segunda etapa fueron, y son, las *viviendas particulares ocupadas por uno o más hogares particulares*, de los cuales se elige al azar *uno*. Las unidades se seleccionan en conglomerados de tres. La selección es aleatoria dentro de la zona.

2. Estratificación de Montevideo

Con la excepción de las muestras del año 1980 y primer semestre de 1981, y las variaciones de cobertura geográfica ya mencionadas, la afijación del número de entrevistas fue siempre *proporcional al tamaño de los estratos* medido igual que en el Interior. Cabe consignar que *nunca* se publicaron estimaciones de parámetros a nivel de estrato.

Las unidades de primera etapa cambiaron de definición a lo largo de la historia de la ECH de Montevideo. Desde 1980 a 1983 y de 1990 a la actualidad, las unidades primarias fueron zonas censales seleccionadas con probabilidad proporcional a su tamaño. Desde 1984 a 1989, y para disminuir el efecto de *correlación intraclase* en los conglomerados de viviendas seleccionadas, las unidades primarias fueron segmentos censales.

Los conglomerados de viviendas tuvieron un tamaño de 5/zona en 1980 y primer semestre de 1981, de 9/segmento cuando las unidades de primera etapa fueron los segmentos, y de 3/zona el resto del tiempo. [5]

Con respecto a la definición de los estratos:

- En 1980 y primer semestre de 1981 consistieron en una división geográfica y socio-económica de zonas en 4 estratos: *alto, medio, bajo* y *rural.*

[4] Las □regiones□ fueron: Litoral norte (Artigas, Salto y Paysandú), Centro norte (Rivera, Tacuarembó y Durazno), Nordeste (Cerro Largo y Treinta y Tres), Litoral sur (Río Negro, Soriano y Flores), Suroeste (Colonia, San José, Canelones y Florida), y Sureste (Rocha, Maldonado y Lavalleja). Las estimaciones a nivel □región□ mostraron ser de una precisión excelente en términos de varianzas, si se hace caso omiso de *qué representan exactamente esas regiones en términos socio-económicos o productivos*. No sucede lo mismo cuando las estimaciones se expresan por □departamento□. En este caso, las diferencias entre tamaños muestrales llega a afectar seriamente los intervalos de confianza sobre los parámetros de interés.

[5] Hasta 1992 no se hicieron estimaciones del efecto de diseño□ (*deff*) sobre las estimaciones de varianza de las principales variables estudiadas. Para más información, ver el Anexo Metodológico en la publicación anual de 1991 y en la Metodología de la Encuesta Continua de Hogares ya mencionada en la nota [4].

- Para el segundo semestre de 1981 y hasta 1983, fueron 5 estratos (*alto, medio alto, medio, bajo,* y *rural*), construidos con la información del ingreso de los hogares recogida en el período anterior.
- En 1984, basándose en información de *ocupación de las personas que predominaba a nivel de segmento* y que fue recogida en el año 1982, se dividió Montevideo en 3 estratos "ocupacionales" A, B (llamados entre casa: "corbatas" y "overoles" respectivamente) y AB (cuando no había predominio de ninguna de esas clases).
- En 1990, y con información del Censo 1985, se clasificaron las *zonas* según los criterios de estratificación de 1984, donde el "predominio" se definió en una proporción mayor o igual a 2/3 entre las ocupaciones clasificadas en A o B.

Unas últimas palabras: hasta 1993, las publicaciones anuales (en papel) llevaron una introducción extensa acerca del diseño y principales definiciones de variables relevadas por la ECH. Una vez que se redactó y publicó la ***Metodología*** ya citada en la nota [4], se eliminó la parte del diseño muestral de la introducción a los informes anuales. Por otra parte, éstos se convirtieron en una presentación comentada de los cuadros o tabulaciones más consultadas, teniendo en cuenta que el resto de las tabulaciones estándar pasaron a ofrecerse en *diskette* a los usuarios que las solicitan.

3. Niveles Geográficos de Estimación

Aunque regularmente se ha planteado al INE que sería de interés contar con datos provenientes de una encuesta a hogares para todo el país incluyendo el área rural, consideraciones de costo y recursos humanos han postergado una y otra vez esas iniciativas. [6] El año 1998 no será la excepción, a menos que el Ministerio de Ganadería, Agricultura y Pesca logre poner en práctica este año el proyecto de levantar una Encuesta de Hogares en el área rural, según criterios de estratificación marcados por el tipo de producción (agropecuaria, agrícola, ganadera, lechera, arrocera y granjera, entre otras).

Por lo tanto, una vez más la ECH será aplicada en Montevideo y en el Interior *urbano*.

Un fenómeno de aparición relativamente reciente y de gran importancia en la desagregación geográfica de estimaciones de variables fundamentales tales como la tasa de desocupación y el ingreso de los hogares, es el que se ha dado en llamar (extraoficialmente) *área metropolitana* (de Montevideo) o *Gran Montevideo*. Esta "región" estaría constituida por el propio departamento de Montevideo, más las localidades aledañas de los departamentos de Canelones y San José, esto es, la *Periferia* de Montevideo [7].

Durante las etapas de preparación, relevamiento y procesamiento de la *Muestra de Anticipación de Resultados Censales* del Censo General de 1985, se reservó una llamada "Área Periférica" que abarcaba *secciones censales* enteras de estos dos departamentos (la 4̃ y 16̃ de Canelones y la 6̃ de San José. Ver en el Apéndice 1 la definición de sección censal).

Desde 1980, la Periferia estuvo representada en parte por las ciudades de Las Piedras y Pando y sólo en las últimas encuestas anuales por algunas otras localidades ubicadas en la actual *Ciudad de la Costa* y periferia de Canelones (por ejemplo, Solymar, La Paz y los fraccionamientos de Camino Maldonado). La periferia localizada en San José no ha sido alcanzada por la ECH.

---

[6] Como se mencionó en el punto III, solo una vez la ECH alcanzó a recoger datos en centros poblados muy pequeños y parajes rurales. De esta experiencia y de otras actividades estadísticas que alcanzaron la cobertura total del país (el censo económico, por ejemplo) se ha calculado que el costo por unidad de información en el área rural es, en cifras gruesas, diez veces el costo por unidad en el área urbana.
Si bien ningún paraje del país es inaccesible (salvo fenómenos pasajeros como inundaciones) y el acceso se va facilitando con la mejora de la caminería y otros medios de comunicación, hay diversos costos de acceso que no pueden eliminarse, entre ellos combustible para vehículos (cuando no hay ómnibus) y viáticos.

[7] La Periferia comprende las comúnmente llamadas ciudades dormitorio habitadas por personas que en proporción importante se desplazan a Montevideo principalmente para trabajar. Qué tan importante es ese desplazamiento y las razones de ello (otras que trabajo) no son investigadas en la ECH, sea porque la batería de preguntas del cuestionario no las consideran (por ejemplo, si se desplazan para atención de salud o adquisición de bienes y servicios), o sea porque el alcance geográfico de la encuesta es insuficiente.

Las ciudades mencionadas y localidades menores de la Periferia de Canelones se agruparon para estimar las variables *de Canelones* y no para identificar el fenómeno del *Gran Montevideo.* [8]

Considerados los pro y contra de los diseños pasados (por ejemplo, agotamiento de zonas en localidades menores, aparte de su más alto costo por unidad de información y baja cantidad de unidades), se decidió que la ECH 1998 tendrá los siguientes niveles geográficos de estimación:

1. La capital de la República.
2. El *Gran Montevideo*, constituido por la capital más la Periferia la cual comprende *todas las localidades urbanas* (amanzanadas) *en un radio de 30 Km de Montevideo* (a contar desde el Km 0).
3. El Interior Urbano, limitado a la estimación de variables en las *localidades amanzanadas mayores*. En este sentido, se redujo el número de localidades *elegibles* a dos grupos:

    a) *Las localidades mayores*. Se llamaron así las localidades de 15.000 o más habitantes según los datos del Censo 1996. Incluye todas las capitales departamentales (18) y otras ocho localidades. Este grupo de centros poblados es de inclusión forzosa. En doce de ellas se extendieron los límites para incluir los barrios "conurbados", cuyas condiciones deberían ser de interés investigar: Artigas, Bella Unión, Canelones, Melo, Durazno, Minas, Maldonado, San Carlos, Punta del Este, Paysandú, Fray Bentos, Rivera y Treinta y Tres.

    b) *Las localidades medianas*. Son las de 5.000 a 15.000 habitantes. Incluye 20 localidades distribuidas en 12 departamentos (Artigas, Flores, Lavalleja, Paysandú, Salto y Treinta y Tres no tienen). Se seleccionó ***una*** localidad mediana por cada departamento que tiene al menos una, con una probabilidad proporcional a su tamaño en número de habitantes. Por lo tanto, son 12 centros poblados seleccionados en este grupo.

El resto urbano del país y las áreas rurales largamente omitidas en las investigaciones continuas (fuera de la de Montevideo), en el futuro, podrían ser merecedoras de la aplicación de relevamientos especiales periódicos para mantener la información al día en esas áreas y no esperar exclusivamente a los censos decenales para tenerla. Además podrían buscarse mecanismos muestrales con los cuales poder efectuar *estimaciones insesgadas* del país en su conjunto y no tener que recurrir a la necesidad de explicar a los usuarios (sean calificados o no) cuales son exactamente los alcances de la cobertura geográfica de la ECH.

4. Localidades del Interior Urbano y de la Periferia

Según el grupo de tamaño al que pertenece una localidad del Interior urbano (no Periferia), habrá tenido que ser sorteada (con probabilidad proporcional a su tamaño medido en cantidad de viviendas particulares ocupadas), o incluida forzosamente.

A continuación se muestra la tabla de localidades incluidas en la muestra de la ECH 1998. En letra *itálica* o *bastardilla* están indicados los centros poblados del grupo mediano, y llevan asterisco (*) las que en ese grupo resultaron seleccionadas cuando el departamento tiene más de una localidad. Estas últimas van acompañadas de la probabilidad con que fueron seleccionadas dentro de su grupo. Las localidades mayores tienen probabilidad **1**.

[8] Si se interpretan con liberalidad e imaginación los signos mercadotécnicos del comercio y del transporte, se puede delimitar sobre un plano, lápiz en mano, la extensión de la Periferia. Por ejemplo, observando los destinos de las compañías de omnibuses interdepartamentales de corto y mediano alcance, aquellos cuya frecuencia de salida supera 2 por hora *son, por lo común, parte de la Periferia.*

Personas, viviendas particulares y número de zonas, según nivel geográfico de estimación y localidad. Datos del Censo'96.

Departamento	Localidad		Total personas	Personas en viv.part.	Viviendas particulares ocupadas	Zonas
TOTAL			*2.651.487*	*2.615.644*	*814.726*	*30.715*
	De 15.000 o más		**2.473.092**	**2.438.997**	**757.276**	**26.383**
	De 5.000 a 15.000		*178.395*	*176.647*	*57.450*	*4.332*
Montevideo			**1.344.839**	**1.321.576**	**424.227**	**8.591**
Interior urbano			**1.306.648**	**1.294.068**	**390.499**	**22.124**
Artigas	Artigas		41.488	41.156	11.383	518
	Bella Unión		16.801	16.708	4.469	283
Canelones	PF - Canelones		191.474	190.508	53.567	2.675
	PF - Ciudad de la Costa		78.576	78.259	23.770	1.510
	Canelones		19.857	19.307	6.051	297
	Santa Lucia		16.764	16.430	5.009	350
	San Ramón (*)	*0,5475*	*6.828*	*6.610*	*2.189*	*162*
	Salinas		*5.279*	*5.275*	*1.809*	*175*
Cerro Largo	Melo		47.788	47.401	14.557	663
	Rio Branco		12.215	12.159	3.919	288
Colonia	Colonia		22.200	21.762	7.180	421
	Carmelo		16.658	16.471	5.103	334
	Nueva Helvecia (*)	*0,2088*	*9.650*	*9.467*	*3.206*	*244*
	Juan Lacaze		*12.988*	*12.926*	*4.271*	*228*
	Rosario		*9.428*	*9.287*	*3.057*	*190*
	Nueva Palmira		*8.339*	*8.262*	*2.771*	*327*
	Tarariras		*6.174*	*6.082*	*2.051*	*115*
Durazno	Durazno		31.850	31.592	8.988	556
	Sarandí del Yi	*1,000*	*6.662*	*6.616*	*2.148*	*232*
Flores	Trinidad		20.031	19.683	6.387	478
Florida	Florida		31.594	31.152	9.216	458
	Sarandí Grande	*1,000*	*5.635*	*5.574*	*1.863*	*146*
Lavalleja	Minas		37.233	36.657	12.079	571
Maldonado	Maldonado		49.747	49.423	15.897	594
	San Carlos		24.168	23.957	7.939	355

Personas, viviendas particulares y número de zonas, según nivel geográfico de estimación y localidad. Datos del Censo'96.

Departamento	Localidad		Total personas	Personas en viv.part.	Viviendas particulares ocupadas	Zonas
	Punta del Este		15.283	15.074	6.648	908
	Piriapolis (*)	*0,5675*	*7.570*	*7.497*	*2.713*	*404*
	Pan de Azúcar		*6.532*	*6.523*	*2.068*	*158*
Paysandú	Paysandú		81.822	80.927	22.784	1.051
Río Negro	Fray Bentos		22.623	22.400	6.322	372
	Young	*1,000*	*14.567*	*14.429*	*3.970*	*276*
Rivera	Rivera		71.345	70.873	20.826	966
	Tranqueras	*1,000*	*5.792*	*5.784*	*1.758*	*202*
Rocha	Rocha		26.017	25.533	9.089	619
	Chuy (*)	*0,3717*	*9.804*	*9.754*	*3.268*	*152*
	Castillos		*7.346*	*7.278*	*2.919*	*197*
	Lascano		*7.134*	*7.076*	*2.604*	*181*
Salto	Salto		93.117	92.030	24.097	1.075
San José	San José de Mayo		34.552	34.111	10.563	508
	PF - San José		20.070	20.022	5.790	478
	Libertad	*1,000*	*8.353*	*8.291*	*2.612*	*138*
Soriano	Mercedes		39.320	38.971	11.447	494
	Dolores	*1,000*	*14.784*	*14.648*	*4.385*	*307*
Tacuarembó	Tacuarembó		45.891	45.408	13.681	611
	Paso de los Toros	*1,000*	*13.315*	*13.109*	*3.869*	*210*
Treinta y Tres	Treinta y Tres		31.984	31.606	10.207	647

La Periferia de Montevideo es tratada como "una ciudad" a la hora de seleccionar las unidades primarias y secundarias.

La Periferia está compuesta de tres sub-áreas que se designan: *Periferia de Canelones*, *Ciudad de la Costa* y *Periferia de San José*. Como ya se indicó, constituyen un "cinturón" de localidades (ciudades, pueblos, villas, fraccionamientos, barrios y otros centros poblados) de diversos tamaños más o menos conurbadas entre sí, que se encuentran entre el límite departamental con Montevideo en un "semicírculo" hasta el Km 30 aproximadamente.

Las localidades de Canelones y Ciudad de la Costa están distribuidas en las secciones censales 4, 5, 6, 16 y 37 de este departamento. Las de San José se encuentran en parte de la sección 6 de este departamento.

La Periferia de Canelones abarca 23 localidades; la Ciudad de la Costa, 10; y la Periferia de San José incluye 4 centros poblados.

5. Períodos de Desagregación

Las crecientes demandas de información por parte de los usuarios han llevado a que la difusión de los resultados de las variables básicas (actividad, empleo, desempleo e ingresos) se produzca mensualmente en la forma de "trimestre móvil" tanto para Montevideo como para el Interior.

Por otra parte, gobiernos departamentales, prensa e instituciones públicas o privadas cuyas actividades alcanzan a las poblaciones del Interior, piden información desagregada por departamentos y, a veces, por localidades.

Estos pedidos no siempre se pueden satisfacer manteniendo al mismo tiempo confiabilidad, oportunidad y actualidad. Por ejemplo, la desagregación por localidades no se hace siquiera sobre una base anual, ya que los tamaños muestrales de los centros poblados más pequeños no son suficientes para soportar aperturas que presenten un razonable coeficiente de variación.

Por razones que se exponen en la sección siguiente (VII), la muestra del Interior ha sido calculada de modo que puedan ofrecerse estimaciones en las variables básicas sobre una base trimestral a nivel de los grandes dominios (Montevideo, Gran Montevideo e Interior urbano), y sobre una base anual a nivel de departamento.

6. La Estratificación Vigente

En la metodología de la estratificación próxima a publicarse, se presenta el ejercicio realizado para actualizar y optimizar los diversos criterios de estratificación que se aplicaron a los sucesivos diseños de la muestra de la ECH a lo largo de su historia, como se ha relatado en la sección III.

A continuación se resumen los resultados de ese ejercicio.

1. Montevideo y Periferia

Como ya se explicó en la sección IV, para dar cuenta del fenómeno reconocido con el nombre (no oficial) de *Área Metropolitana de Montevideo*, se extendieron los límites departamentales de Montevideo hasta abarcar todas las localidades urbanas de Canelones y San José en un radio de unos 30 Km a partir del Km 0.

Las localidades pertenecientes a las jurisdicciones de estos dos departamentos se dividen en tres, como ya se mencionó en la página 10: Ciudad de la Costa, Periferia de Canelones, y Periferia de San José, esta última incorporada a la muestra a partir de la "ronda" de 1998.

La Periferia como un todo constituye, al menos durante 1998, *un estrato del Gran Montevideo*. A pesar que las características de las localidades ya residentes en la ECH reflejan un fuerte parecido al segundo estrato de Montevideo, se prefiere mantener esta área separada porque: (a) esas localidades no la cubren uniformemente (la Periferia de San José nunca tuvo representación) y en consecuencia (b) no se conoce aun su "estructura" en cuanto a la variable de segmentación utilizada para Montevideo.

Montevideo, por su parte, ha quedado segmentado en cuatro estratos socio-económicos, a saber: *bajo, medio bajo, medio alto* y *alto*.

La única variable de segmentación utilizada fue *el ingreso medio per cápita de los hogares*, a nivel de *segmento censal*, expresado en UR (unidades reajustables). En Montevideo, esta variable está altamente relacionada con la tasa de desempleo (negativamente).

Véase en la siguiente tabla los resultados del último año de encuesta según la clasificación en cuatro estratos de Montevideo:

ECH 1997. Ingreso medio per cápita, tasa de desocupación y margen de confianza al 95% según trimestre y estrato de Montevideo

Trimestre	Estrato	Ingreso medio pc (UR)	. t s_{ing}	Tasa de desocupación	. t s_{des}
	Mvd-bajo	23,69	2,88	13,51	2,72
	Mvd-medio bajo	22,21	1,43	11,94	2,06
	Mvd-medio alto	30,24	1,74	10,87	1,88
1er Trimestre	Mvd-alto	45,26	4,02	10,54	2,16
	Mvd-bajo	21,79	2,65	14,03	2,94
	Mvd-medio bajo	19,76	1,29	15,59	2,55
	Mvd-medio alto	28,82	1,80	12,63	1,94
2ı Trimestre	Mvd-alto	46,44	3,94	9,04	2,29
	Mvd-bajo	20,77	2,43	16,28	3,31
	Mvd-medio bajo	19,71	1,37	16,01	2,55
	Mvd-medio alto	29,74	1,76	10,35	1,80
3er Trimestre	Mvd-alto	50,50	4,00	8,02	1,88
	Mvd-bajo	20,41	2,14	9,77	2,20
	Mvd-medio bajo	19,99	1,18	11,81	2,14
	Mvd-medio alto	29,36	1,76	9,39	1,80
4ı Trimestre	Mvd-alto	47,47	3,55	8,29	1,80

Los estratos *bajo* y *medio bajo* parecen constituir un solo grupo considerando la variable ingreso *per cápita.* Esto sugiere que la segmentación de estos dos estratos podría mejorarse con la incorporación de más información. Este ejercicio se hará cuando se haya procesado la muestra de 1998, la cual, al cubrir la población de hogares con un marco muestral más actualizado puede permitir la evaluación tanto de la estratificación de los segmentos que fueron sorteados como la eficiencia del diseño.

2. Resto urbano del Interior

La misma metodología que permitió la segmentación de Montevideo se aplicó a las localidades departamentales que representan el área urbana del Interior para obtener una estratificación regional de este territorio. Para conocer cuál fue el agrupamiento y cómo se le designó el lector puede consultar el mismo documento ya mencionado.

En la sección VI se mencionó que la principal razón por la cual se difunden resultados de las áreas urbanas departamentales es la demanda de los usuarios. Por lo tanto, a pesar que se habría obtenido una disminución sustancial de las cargas departamentales utilizando la estratificación, se tomó una decisión estratégica en el sentido de prever la producción de estimaciones por departamento con cierta regularidad.

En 1996, la dispersión de las estimaciones de la tasa de desocupación entre departamentos fue bastante importante. Esto solo hizo que la mejora lograda por la reducción del número de localidades urbanas alcanzadas por la muestra, y por la estratificación de Montevideo con la Periferia pasase casi inadvertida. De hecho, la muestra mensual en 1998 es un 24% menor a la muestra mensual de 1997. Teniendo en cuenta que las muestras de 1997 y anteriores estaban sobredimensionadas en un 20% para absorber los problemas de marco, la ganancia neta en 1998 es de solo un 4%.

7. Tamaño de la Muestra Anual

El tipo de diseño elegido para la selección de las unidades a ser encuestadas tiene el nombre de ***muestra estratificada bietápica donde la unidad primaria de selección es la zona censal y la unidad secundaria es el hogar particular***. En el Interior urbano, debería hablarse en realidad de muestreo ***trietápico***, ya que en unos pocos departamentos hubo que seleccionar una de dos o más localidades con 5.000 a 15.000 habitantes. En este caso, la unidad primaria sería la localidad, la secundaria sería la zona y, por último, la terciaria sería el hogar particular.

Las variables que han sido, y son, determinantes del cálculo de tamaños muestrales apropiados para cada estrato y dominio de estudio son: la ***tasa de desocupación*** y el ***ingreso total medio de los hogares***. Los datos de referencia en los que se basó el cálculo fueron los proporcionados por la ronda 1996 de la ECH.

Como se mencionó antes, la tasa de desocupación presenta una gran dispersión entre departamentos en el Interior urbano (al menos durante 1996) comparada con el ingreso promedio respectivo. Lo opuesto ocurre entre los estratos de Montevideo, agregándose el fenómeno observado de la alta correlación (negativa) entre ingresos y desocupación. [9]

El cálculo de los tamaños muestrales de cada estrato, se tuvo que ajustar a un "compromiso" entre el óptimo para estimar el ingreso, el óptimo para estimar la desocupación y la que daría una simple afijación proporcional según el tamaño del estrato.

1. Muestra de Montevideo y Periferia

Se determinaron los tamaños muestrales por estrato de tal modo que se pudiese estimar ***trimestralmente*** la tasa de desocupación y el ingreso medio *per cápita* de los hogares.

Para la tasa de desocupación se fijó un error absoluto de un 1% de la máxima tasa trimestral y una confianza de 95%. Esto daría un coeficiente de variación algo menor al 5% para esta variable.

Para el ingreso, se fijó un error absoluto del orden del 4% (1 UR : a precios de marzo 1998 son unos 17 dólares) del promedio de los máximos ingresos *per cápita* de los estratos y con una confianza también del 95%. Esto daría, para esta variable, un coeficiente de variación esperado de un 2%.

La *afijación óptima* del tamaño total de la muestra a los estratos se hizo considerando el costo por unidad fijo y distribuyéndola proporcionalmente a la varianza de la variable en el estrato.

Este procedimiento, como era de esperar, dio tamaños muestrales diferentes para cada variable y requerimientos. En estos casos, se recomienda que se busque un "*tamaño compromiso*" entre los tamaños específicos de las variables de interés, y el tamaño que resultaría de calcular un tamaño muestral con requerimientos promedio y afijación proporcional al tamaño de los estratos. Este "compromiso" es el promedio por estrato de los tres tamaños determinados por los tres métodos.

Este método se aplicó no solamente para el Gran Montevideo sino también para el resto urbano del país. Los extractos de las planillas de cálculo se muestran en la sub-sección siguiente.

El coeficiente de variación esperado por trimestre con la aplicación del tamaño promedio sería aun menor al 10% para la estimación de la tasa de desocupación y de un 4% para el ingreso.

Las ecuaciones aplicadas en la planilla de cálculo son las siguientes [10] :

□ Para el ingreso *per cápita* :

[9] En efecto, con los datos recolectados en 1997 y post-clasificados por futuro estrato socio-económico□, se calculó un $r = -0,74$ $(p<0,001)$ (r de Pearson) entre las tasas de desocupación trimestrales y los ingresos (*per cápita*) trimestrales. Si se incorporaba la Periferia, el coeficiente se reducía a $r = -0,72$ $(p<0,001)$.

[10] Extraídas de **Cochran, W.G.** (1977). *Sampling Techniques*. John Wiley & Sons, New York.

$$n = \frac{\left(\sum_{h}^{L} W_h \sigma_h\right)^2}{V + \frac{1}{N}\sum_{h} toL W_h \sigma_h^2} \bullet \frac{d}{c}$$

,

donde σ_h^2 es la varianza de esa variable en el estrato h-ésimo, $W_h = \frac{N_h}{N}$, N y N_h son las cantidades de viviendas particulares ocupadas en la población y en el estrato h-ésimo respectivamente, c es la cantidad media de hogares por vivienda particular ocupada y d es el factor de corrección debido a que, por razones principalmente operativas, se seleccionan grupos de 3 hogares por zona.

Luego, la parte correspondiente de esta muestra en el h-ésimo estrato se calcula :

$$n_h = n \frac{W_h \sigma_h}{\sum_{h}^{L} W_h \sigma_h} = n \frac{N_h \, sigmasubh}{\sum_{h}^{L} N_h \sigma_h}$$

,

considerando que el costo por unidad de información (el hogar) es fijo.

□ Para la tasa de desocupación p, [11] se calcula el tamaño n igual que antes, sustituyendo σ_h^2 por $P_h Q_h$ y $\frac{d}{c}$ por la fracción $\frac{d}{a\,b\,c}$. Las cantidades d y c tienen el mismo significado anterior (d correspondiendo a la variable p) y además a es la proporción de activos y b la cantidad de personas por hogar.

Luego, la afijación para el h-ésimo estrato se calcula:

$$n_h _ n \frac{N_h \sqrt{P_h Q_h}}{\sum_{h}^{L} N_h \sqrt{P_h Q_h}}$$

2. Muestra del resto urbano

Se determinó el tamaño muestral total y por departamento con los mismos requerimientos de error absoluto y nivel de confianza que para el Gran Montevideo, pero aplicados a la estimación ***semestral*** de las principales variables.

A continuación se muestra un extracto de la planilla de cálculo [12] construida para determinar los tamaños muestrales de Montevideo, Periferia y resto urbano para la ECH de 1998. Los "ajustes prácticos" aplicados

[11] Si se es purista, la tasa de desocupación es una razón entre la cantidad de desocupados (variable) y la cantidad de personas económicamente activas (variable). Sin embargo, para un determinado período de referencia y dominio se asume que es una proporción P fija.

[12] Desarrollado en planilla electrónica Quattro Pro v.6.02, 1996 □ Corel Corp.Ltd.

para redondear la cantidad de hogares por estrato se basaron en múltiplos de la cantidad de hogares por zona y en el período de referencia.

Planilla de cálculo de los tamaños muestrales por "opción" y variable, según estrato. Datos de base, 1996.

ESTRATO	Nh	Tamaño Comp. (1)	Tamaño Prop.	Tamaño P	Tamaño ing.pc.	T.Desoc. %	Ing.pc. /sem (2)
Gran MVD + Dptos	**823.079**						
Total Gran MVD	*507.354*	*2.781*	*2.106*	*3.564*	*2.556*	*14,02%*	*28,3*
MVD	*424.227*	*2.205*	*1.755*	*2.947*	*1.823*	*13,58%*	*23,3*
Mvd-bajo	86.190	450	360	670	300	17,58%	18,8
Mvd-medio bajo	100.968	531	423	747	413	15,57%	22,1
Mvd-medio alto	133.858	666	558	903	514	12,47%	20,7
Mvd-alto	103.211	558	432	626	596	9,79%	31,2
Periferia	83.127	576	351	617	733	15,70%	47,6
Total INT.URB.	*315.725*	*3.222*	*2.376*	*3.995*	*2.786*	*11,27%*	*17,3*
ART	15.852	198	119	234	210	15,42%	17,7
CAN (*)	18.748	252	141	259	315	13,12%	30,7
CLA	18.476	180	139	231	151	10,47%	12,9
COL	27.639	252	208	336	199	9,81%	17,8
DUR	11.136	108	84	140	93	10,53%	15,1
FLO	6.387	72	48	67	50	6,98%	19,8
FLA	11.079	90	83	127	58	8,61%	13,8
LAV	13.768	126	104	172	73	10,42%	14,8
MLD	35.265	342	265	466	265	11,83%	17,4
PAY	24.252	234	183	313	157	11,19%	16,9
RNE	10.292	126	77	182	81	24,84%	18,6
RIV	22.584	234	170	302	195	12,12%	16,0
ROC	17.880	198	135	219	192	9,95%	15,4
SAL	24.097	198	181	216	179	5,06%	14,6
SAJ (*)	13.175	144	99	144	159	7,74%	18,2
SOR	17.338	180	130	221	145	10,85%	14,8
TAC	17.550	180	132	227	169	11,22%	21,1
TYT	10.207	108	77	139	95	12,77%	16,7

El tamaño "compromiso" provee una estimación un poco mejor de la variable de menor varianza en cada estrato. Por ejemplo, el ingreso *per cápita* medio en cualquier departamento del Interior (con la posible excepción de Canelones y San José, si no contamos con la Periferia), va a ser estimada con un poco más de precisión que el desempleo, si los niveles de desempleo permaneciesen en los mismos guarismos que los estimados durante 1996. [13]

El tamaño muestral anual en cantidad de viviendas y en cantidad de zonas (a razón de tres viviendas por zona) se muestra en la tabla siguiente:

[13] De hecho, en 1997 la mitad de las tasas de desocupación departamentales han decrecido, y también la tasa de desocupación del Interior urbano, con lo cual se estaría estimando con mejor precisión esta variable. Por otra parte, el ingreso *per cápita* medio no ha variado, en general, en términos de la moneda utilizada, la UR.

Número de viviendas en la muestra por período de referencia, por mes y zonas por mes, según estrato.

ESTRATO	Tamaño Ajustado (1)	Tamaño por mes	Zonas por mes
Gran MVD + Dptos		***1.464***	***399***
Gran MVD	*2.781*	*927*	*309*
MVD	*2.205*	*735*	*245*
Mvd-bajo	450	150	50
Mvd-medio bajo	531	177	59
Mvd-medio alto	666	222	74
Mvd-alto	558	186	62
Periferia	576	192	64
TOTAL DEPTOS	***3.222***	***537***	***179***
ART	198	33	11
CAN (*)	252	42	14
CLA	180	30	10
COL	252	42	14
DUR	108	18	6
FLO	72	12	4
FLA	90	15	5
LAV	126	21	7
MLD	342	57	19
PAY	234	39	13
RNE	126	21	7
RIV	234	39	13
ROC	198	33	11
SAL	198	33	11
SAJ (*)	144	24	8
SOR	180	30	10
TAC	180	30	10
TYT	108	18	6

(1) Gran Montevideo por trimestre, Interior Urbano por semestre.
(*) Excluye Periferia.

El tamaño muestral mensual aplicado hasta 1997 era de 2.000 hogares, la mitad de ellos en Montevideo, en donde se habían verificado los mayores cambios en el marco muestral utilizado.

3. Probabilidad de selección de las unidades; factores de expansión anuales

En función de los estratos definidos y las localidades seleccionadas, se han calculado las probabilidades de selección de las viviendas dentro de localidades según el estrato de tamaño al cual pertenecen, para determinar los factores de expansión que permitirán estimar totales de personas en características tales como actividad económica, empleo o desempleo.

En la página siguiente se muestra la tabla de factores de expansión anuales y probabilidades de selección de las viviendas por grupo de tamaño y localidad que se aplicarán en dichas estimaciones.

PROBABILIDADES DE SELECCIÓN Y FACTORES DE EXPANSIÓN

ESTRATO	LOCALIDAD	Factor de expansión anual	PROBABILIDAD DE SELECCIÓN		ZONAS	
			Vivienda en I	Vivienda en II	Grupo I	Grupo II
Mvd bajo		46,94	0,0213		612	
Mvd medio bajo		46,74	0,0214		720	
Mvd medio alto		49,58	0,0202		900	
Mvd alto		45,57	0,0219		755	
Periferia						
Canelones		35,93	0,0278		497	
Ciudad de la costa		36,18	0,0276		219	
San José		37,12	0,0269		52	
Artigas	Artigas	35,13	0,0285		108	
	Bella Unión	41,38		0,0242		36
Canelones (*)	Canelones	28,01	0,0357		72	
	Santa Lucía	19,88	0,0503		84	
	San Ramón	55,56		0,0180		24
Cerro Largo	Melo	50,51	0,0198		96	
	Río Branco	36,29		0,0276		36
Colonia	Colonia	24,93	0,0401		96	
	Carmelo	47,25	0,0212		36	
	Nueva Helvecia	106,64		0,0094		48
Durazno	Durazno	41,61	0,0240		72	
	Sarandí del Yi	59,67		0,0168		12
Flores	Trinidad	44,35	0,0225		48	
Florida	Florida	51,20	0,0195		60	
	Sarandí Grande	51,75		0,0193		12
Lavalleja	Minas	41,94	0,0238		96	
Maldonado	Maldonado	55,20	0,0181		96	
	San Carlos	44,11	0,0227		60	
	Punta del Este	46,17	0,0217		48	
	Piriápolis	44,27		0,0226		36
Paysandú	Paysandú	45,21	0,0221		168	
Río Negro	Fray Bentos	35,12	0,0285		60	
	Young	36,76		0,0272		36
Rivera	Rivera	48,21	0,0207		144	

PROBABILIDADES DE SELECCIÓN Y FACTORES DE EXPANSIÓN

ESTRATO	LOCALIDAD	Factor de expansión anual	PROBABILIDAD DE SELECCIÓN		ZONAS	
			Vivienda en I	Vivienda en II	Grupo I	Grupo II
	Tranqueras	24,42		0,0410		24
Rocha	Rocha	25,25	0,0396		120	
	Chuy	122,10		0,0082		24
Salto	Salto	55,78	0,0179		144	
San José (*)	San José	41,92	0,0239		84	
	Libertad	72,56		0,0138		12
Soriano	Mercedes	45,42	0,0220		84	
	Dolores	30,45		0,0328		48
Tacuarembó	Tacuarembó	42,23	0,0237		108	
	Paso de los Toros	53,74		0,0186		24
Treinta y Tres	Treinta y Tres	40,50	0,0247		84	

(*) No incluye la Periferia.

Sería conveniente que estos factores de expansión sean revisados anualmente en función de la actualización permanente de los marcos a utilizar, esto es, del mantenimiento de marcos muestrales "vivos" y de alta calidad.

ANEXO 1 : División Geográfica y Estadística del País

Departamento: Coincide con la división político-administrativa del país. Hay 19 departamentos que, a los efectos estadísticos se numeran según la ordenación alfabética de los departamentos del Interior (02 a 19). Montevideo tiene el número 01.

Sección Censal: Es la máxima subdivisión dentro del departamento. Coincidieron con las Secciones Judiciales al relevarse el IV Censo General de Población y II de viviendas del 16 de octubre de 1963.

Desde entonces, los límites seccionales originales han permanecido sin cambios, produciéndose solamente subdivisiones con efectos temporales limitados para facilitar el relevamiento censal o para dividir (partir) secciones muy pobladas o extensas. Por ejemplo, la sección 1$\tilde{1}$ de Montevideo fue partida en dos luego del Censo 1963: la 1$\tilde{1}$ propiamente dicha y la 9$\tilde{9}$, codificación que las identifica hasta el presente. La actual 1$\tilde{0}$ fue dividida en tres: 10A, 10B y 10C para el levantamiento de los censos de 1985 y 1996 con efectos solamente temporales. Otro ejemplo es la $\tilde{7}$ de Canelones, que fue dividida permanentemente en tres: 1$\tilde{7}$, 2$\tilde{7}$ y 3$\tilde{7}$.

Actualmente hay 231 secciones censales en el país. En los departamentos la cantidad varía entre 6 (Flores) y 26 (Montevideo).

Las secciones censales se subdividen en segmentos censales.

Segmento Censal: Es una subdivisión dentro de cada Sección Censal. Limita un grupo de las llamadas *zonas censales*. Por sección censal hay, en promedio, 17 segmentos. El rango máximo varía entre 1 y 125 segmentos por sección (el máximo ocurre en la sección 1$\tilde{0}$ de Montevideo). Actualmente hay 3.961 segmentos censales en todo el país.

Un segmento está constituido, en promedio, por unas 12 zonas. El rango máximo varía entre 1 y 54 zonas/segmento (el máximo ocurre en la sección $\tilde{9}$, segmento 10 de Rivera).

En general, los segmentos pertenecen a una sola área, esto es, hay segmentos censales urbanos ó segmentos censales rurales pero no mezclas, salvo cuando las localidades urbanas crecen a expensas de las zonas rurales. En este caso, se acostumbra a "re-segmentar" el territorio alterado, sea partiendo los segmentos afectados o bien cambiando el carácter del segmento en cuestión (por ejemplo, todo un segmento rural pasa a ser urbano).

Para el Censo 1996 segmentos excesivamente grandes en términos de la cantidad de viviendas comprendidas, fueron partidos permanentemente para facilitar el control de campo del personal que dirige el trabajo de los empadronadores (los jefes de segmento).

Zona censal: Es la menor subdivisión del "territorio censal". En el área urbana, una zona corresponde generalmente a una *manzana* (en inglés *block*). En el área rural está determinada por límites físicos de fácil reconocimiento en el terreno. Actualmente hay mas de 46.000 zonas en todo el país. El número medio de zonas por sección es de unas 200. Por departamento el número de zonas varía entre 636 (Flores) y 8.860 (Canelones).

Localidad (urbana): Es toda entidad de población de carácter *urbano*. Al Censo 1996 se consideraron como localidades urbanas aquellas que correspondían a centros poblados *amanzanados* que ocupaban la superficie de al menos un segmento censal.

La definición utilizada por el INE en cuanto a la delimitación de áreas urbana-rural se basan en criterios prácticos y de tipo operativo, y parcialmente en las disposiciones de la Ley N° 10.723 del 21/04/46 llamada "Ley de Centros Poblados" y sus modificaciones posteriores. En ella se declara de competencia exclusiva de los gobiernos departamentales la autorización de subdividir predios rurales con destino a la formación de centros poblados, así como para aprobar el trazado y apertura de calles, caminos o cualquier otra vía de tránsito que indique o no amanzanamiento o formación de dichos

centros. También se establecen las superficies de los predios dentro de las zonas urbanas, suburbanas y rurales, y se fijan los requisitos que la Ley califica como mínimos: la existencia de agua potable, las condiciones del terreno y área contigua, y otros servicios indispensables.

Sin embargo, los cambios introducidos periódicamente con el fin de perfeccionar las definiciones censales de referencia, dificultan las comparaciones directas para los diversos fines con que es requerida la información censal.

MUESTRA MAESTRA CENSO 1990 V/S MUESTRA MAESTRA CENSO 2001

INSTITUTO NACIONAL DE ESTADÍSTICA
VENEZUELA

ÍNDICE

Página

Aspectos Generales 321
Marco de Muestreo 321
Diseño Muestral de la EHM 322
Otras Investigaciones 323

MUESTRA MAESTRA CENSO 1990 VS. MUESTRA MAESTRA CENSO 2001. PAIS: VENEZUELA

ASPECTOS GENERALES

	CENSO 1990	CENSO 2001
Fecha de levantamiento	Octubre de 1990	Octubre-Noviembre 2001
Fecha conformación Muestra Maestra	1994	Tercer trimestre 2001
Fecha de la primera aplicación a la EHM	Segundo semestre 1997	Segundo semestre 2001. (Información pre-censal)

MARCO DE MUESTREO

	MUESTRA MAESTRA CENSO'90	MUESTRA MAESTRA CENSO'01
Descripción	Marco Mixto: *Base de datos de segmentos y sectores *Carpetas de segmento/sector con las viviendas	Marco Mixto: *Base de datos de segmentos y sectores *Carpetas de segmento/sector con las viviendas
Materiales de la carpeta	*Planilla de selección de áreas *Planilla de identificación de áreas *Listado de viviendas para cada área sel. *Croquis del área seleccionada *Plano del segmento ó del sector	*Planilla de selección de Lotes *Planilla de identificación del lote seleccionado *Listado de viviendas para cada lote sel. *Croquis del lote seleccionado *Plano del segmento ó del sector
Variables auxiliares que contiene el marco	*Mapa de pobreza *Urbano/rural *Barrio-No barrio *Estrato *Dominio (Entidad-AM-Resto) *Dominio (AMC, Ciu. Grandes, Ciu. Medianas, Ciu. pequeñas) *Total de viviendas en el segmento/sector	*Dominio (Entidad-AM-Resto) *Dominio (AMC, Ciu. Grandes, Ciu. Medianas, Ciu. pequeñas) *Total de viviendas en el segmento/sector. NOTA: se ha planificado realizar post-estratificación una vez que se la información del censo 2001 este transcrita
Proceso de actualización	La actualización se realiza en campo con el material cartográfico correspondiente al censo	*La actualización se realiza en oficina con el registro de estructura y carpeta parroquial rural. *Las encuestas a ser lev. muy distante a la fecha del lev. censal su act. se hará en campo
Unidades actualizadas	*La actualización de segmentos y áreas no seleccionadas por la EHM se realiza con la planificación de una encuesta. *La actualización de las áreas seleccionadas por la EHM es semestral con el levantamiento continuo de la EHM	*La actualización de segmentos y lotes no seleccionados por la EHM se realiza con la planificación de una encuesta. *La actualización del lote seleccionado por la EHM es semestral con el levantamiento continuo de la EHM
Tratamiento de viviendas adicionales, omitidas, etc.	*El encuestador actualiza el listado de vivienda como lo encuentra en el terreno. Añade las viviendas nuevas y tacha las desaparecidas.	*El encuestador actualiza el listado de vivienda como lo encuentra en el terreno. Añade las viviendas nuevas y tacha las desaparecidas.

DISEÑO MUESTRAL DE LA EHM

Marco de muestreo	Censo 1990. Estratificado con PPT de viviendas	Censo 2001. Estratificado con PPT de viviendas
Clase de muestreo	Trietápico bifásico estratificado sistemático	Bietápico estratificado aleatorio
Unidades de muestreo		
Unidad de primera etapa (UPM)	Segmentos y sectores del censo'90. (I - II FASE)	Segmentos y sectores del censo'01
Definición	**Segmentos:** áreas geograficas de 200 viv. **Sectores:** áreas geograficas de 100 viv.	**Segmentos:** áreas geograficas de 200 viv. **Sectores:** áreas geograficas de 100 viv.
Método de selección	Selección con prob. proporcional al no. de viv. del censo 1990	Selección con prob. proporcional al no. de viv. del censo 2001
Total de seg/sec en el censo	19.090 Segmentos + 4.511 Sectores =23.601	28.201 Segmentos+ 4.766 Sectores =32.967
Total de UPM del censo seleccionadas en la EHM	I FASE: 5783 Segmentos+ 712 Sectores = 6499 II FASE: 2346 Segmentos + 268 Sectores = 2614	3.992 Segmentos + 402 Sectores =4.394
Unidad de segunda etapa (USM)	Areas	Lotes
Definición	Areas: 50 viv en promedio	Lotes:15 viv en promedio
Método de selección	Selección de áreas con probabilidad igual	Selección de lotes con probabilidad igual
Total de USM dentro de las UPM	Segmento: Aproximadamente 4 áreas Sector: Aproximadamente 3 áreas	Segmentos: Aproximadamente 13 lotes. Sector: Aproximadamente 6 lotes
Total de USM seleccionadas dentro UPM	*Dos (2) áreas en cada segmento. *Un (1) área en cada sector	Un (1) lote en cada segmento y sector
Unidad de tercera etapa (UTM)	Viviendas	-
Definición	Estructura destinada al alojamiento de personas	-
Método de selección	Selección sistemática de viviendas con probabilidad proporcional al total de viviendas	-
Total de UTM dentro de las USM	50 Viviendas en cada área	-
Total de UTM seleccionadas dentro USM	5 viviendas por área en segmentos y 10 viviendas por área en los sectores	-
Estratificación	Areas metropolitanas y resto, dentro de cada entidad federal	Entidad Federal
Dominios de Estudio	*Nacional *Area Metropolitana de Caracas * Zulia *Lara *Portuguesa *Tachira *Aragua	Entidad Federal
Ventajas del muestreo utilizado	*Menor error de muestreo con un tamaño de muestra menor *USM actualizada para otras investigaciones	*Muestra concentrada *Menor recorrido para el levantamiento *Disminución de la no-respuesta
Factor de expansión de personas	Estimador de razón	Estimador de razón
Tipo de estimador	*Resultado mensual: Estimador combinado *Resultado trimestral: Estimador separado *Resultado semestral: Estimador separado	*Resultado mensual: Estimador combinado *Resultado trimestral: Estimador separado *Resultado semestral: Estimador separado
Variable auxiliar	Proyecciones de población desagregada por: *Grupos quinquenales de edad *Sexo	Proyecciones de población desagregada por: *Grupos quinquenales de edad *Sexo
Razones de utilización de un nuevo marco	*Actualización de la muestra, el marco utilizado para EHM correspondia al censo 80.	*Ruptura del mallado del censo 1990 en el censo 2001 *Actualización cartográfica reciente del censo 2001
NOTA	Para el período 1994-2000 se diseño una Muestra Maestra. La EHM es una submuestra de la Muestra Maestra	Para el período 2001-2010 la EHM conforma la Muestra Maestra.

OTRAS INVESTIGACIONES		
Encuestas realizadas y/o por realizar.	Encuesta de Hogares por muestreo (EHM). 1997-2001	Encuesta de Hogares por muestreo (EHM). 2001-2002
	Encuesta de Presupuestos Familiares (ENPF). 1997-1998	Encuesta Continua de Presupuestos Familiares (ECPF). A ser ejecutada en el 2003
	Encuesta de Victimización. 1997	Encuesta Social (ENSO). A ser ejecutada en el 2003-2006.
	Encuesta de Población y Familia (ENPOFAM). 1998	Encuesta de Informalidad Fase 2. A ser ejecutada en el 2003.
	Encuesta Social (ENSO). 1998	Encuesta de Informalidad Fase 3. A ser ejecutada en el 2003.
	Encuesta de Conglomerados de indicadores multiples (MICS). 1999	
	Encuesta de Informalidad Fase 1 y 2. La Fase 1 se tomó de la EHMy para la fase 2 se realizó una piloto en el Area Metropolitana de Caracas.	

PROGRAMA PARA EL MEJORAMIENTO DE LAS ENCUESTAS Y LA MEDICION DE LAS CONDICIONES DE VIDA EN AMERICA LATINA Y EL CARIBE (MECOVI)
BID - BANCO MUNDIAL - CEPAL

9° TALLER REGIONAL
DISEÑO Y CONSTRUCCIÓN DE LOS MARCOS DE MUESTREO PARA LAS ENCUESTAS DE HOGARES
Lima, Perú, 17 al 19 de junio de 2002

INEI **CEPAL**

Lista de Participantes

ARGENTINA

CECILIA DAMONTE
Coordinadora de los Marcos Muéstrales para Encuestas a Hogares
Instituto Nacional de Estadística y Censos (INDEC)
Pte. Julio A. Roca 609,
Buenos Aires, Argentina
Tel.: (54-11) 4349 9236
Fax: (54-11) 4349 9886
E-mail: cdamo@indec.mecon.gov.ar

MARTA RODRIGUEZ DE MESSERE
Jefe del Departamento de Muestreo – Encuesta Permanente de Hogares
Instituto Nacional de Estadística y Censos (INDEC)
Pte. Julio A. Roca 609,
Buenos Aires, Argentina
Tel.: (54-11) 4349 9278
Fax: (54-11) 4349 9218
E-mail: mmess@indec.mecon.gov.ar

BOLIVIA

FERNANDO RIVERO SUGUIURA
Responsable del Area de Muestreo-Mecovi
Instituto Nacional de Estadística
Calle Carrasco 1391
La Paz, Bolivia
Tel.: (591-2) 2222333
Fax: (591-2) 2222885
E-mail: mecovi_02@ine.org.bo; fors20@latinmail.com

JUAN FERNANDO MOLLINEDO VALDEZ
Estadístico - Diseño de Muestras Maestras
Instituto Nacional de Estadística
Calle Carrasco 1391
La Paz, Bolivia
Tel.: (591-2) 2222333
Fax: (591-1) 2222885
E-mail: fermi@ceibo.entelnet.bo

CHILE

ARTURO LEIVA GUTIERREZ
Coordinador Nuevo Marco de Áreas para Encuestas de Hogares
Instituto Nacional de Estadísticas
Avenida Bulnes 418
Santiago - Chile
Tel.: (56-2) 4708144
Fax: (56-2)
E-mail: arturo.leiva@ine.cl

FIDELINA URIBE REYES
Estadístico en Depto. Estadísticas de Hogares
Instituto Nacional de Estadísticas
Avenida Bulnes 418
Santiago - Chile
Tel.: (56-2) 4708135
Fax: (56-2)
E-mail: fidelina.uribe@ine.cl

COLOMBIA

SERGIO ENRIQUE ACOSTA MORENO
Coordinador Equipo de Diseños Muéstrales
Departamento Administrativo Nacional de Estadística
Transversal 45 # 26-70, Interior 1 CAN, Edificio DANE
Bogotá, Colombia
Tel.: (571) 5978300 ext. 2313/2254
Fax: (571) 5978330
E-mail: seacostam@dane.gov.co

NELLY SOFIA DUQUE MUÑOZ
Profesional Encargada del Manejo de la Muestra Maestra y de la Encuesta Continua de Hogares
Departamento Administrativo Nacional de Estadística
Transversal 45 # 26-70, Interior 1 CAN, Edificio DANE
Bogotá, Colombia
Tel.: (571) 5978300 ext. 2582
Fax: (571) 5978330
E-mail: nduque10@latinmail.com; nsduquem@dane.gov.co

COSTA RICA

GISELLE ARGÜELLO VENEGAS
Coordinadora del Equipo de Trabajo de Muestreo
Instituto Nacional de Estadística y Censos
Costado Sur del Parque Central, Avenida 4
San José , Costa Rica
Tel.: (506) 2230815
Fax: (506) 2230813
E-mail: gisellea@inec.go.cr

CUBA

ALINA CONCEPCION ALFONSO LEON
Investigador Agregado
Oficina Nacional de Estadísticas
Paseo #60, esquina 5ta. Vedado
Ciudad Habana, Cuba
Tel.: (537) 8324855, 8305021 ext. 386
Fax: (537) 333083
E-mail: leon@one.gov.cu

ECUADOR

MARIA ESTHER ESPINEL
Analista de Investigación y Proyectos Estadísticos 3
Instituto Nacional de Estadística y Censos (INEC)
Juan Larrea N15-36 y José Riofrío
Quito, Ecuador
Tel.: (593-2) 2234372
Fax: (593-2) 2509836
E-mail: inec1@ecnet.ec; rocioespinel@hotmail.com

EL SALVADOR

DANIEL FLORES DE PAZ
Muestrista
Dirección General de Estadística y Censos
Av. Juan Bertis #79
Delgado, El Salvador
Tel.: (503) 2765900
Fax: (503) 2862505
E-mail: mecoviels@salnet.net; dfloresdepaz@yahoo.com

GUATEMALA

ARMANDO CONTRERAS DE LA ROCA
Analista Programador
Instituto Nacional de Estadística
8va. Calle 9-55 Edificio América, Zona 1
Guatemala, Guatemala
Tel.: (502) 2300728
Fax: (502) 2320970
E-mail: mecovi@concyt.gob.gt

HONDURAS

PAULETTE ZÚNIGA FLORES
Gerente de Censos y Encuestas
Instituto Nacional de Estadística (INE)
Blvd. Suyapa, Col. Florencia
Tegucigalpa, Honduras
Tel.: (504) 2398768 - 2397161
Fax: (504) 2398612
E-mail: paulettezuniga@yahoo.com

HORACIO LOVO PERALTA
Técnico de Metodología
Instituto Nacional de Estadística (INE)
Blvd. Suyapa, Col. Florencia
Tegucigalpa, Honduras
Tel.: (504) 2397161
Fax: (504) 2398612
E-mail: love_ho@yahoo.com

NICARAGUA

KARLA ARRIOLA
Responsable del Departamento de Muestras
Instituto Nacional de Estadísticas y Censos
Frente al Hospital Antonio Lenin Fonseca
Managua, Nicaragua
Tel.: (505-2) 2681791, 2681788
Fax: (505-2) 2662031 / 2681792
E-mail: karriola@inec.gob.ni

MEXICO

JAIME MOJICA CUEVAS
Subdirector de Diseño Muestral
Instituto Nacional de Estadística, Geografía e Informática
Rep. Mexicana, Esq. República del Salvador s/n,
Fracc. El Dorado
Aguascalientes, México
Tel.: (52-44) 910 5739
Fax: (52-44) 978 7959
E-mail: Mojica_Jaime@yahoo.com

PANAMA

ROBERTO CASTILLO G.
Jefe de la Sección de Muestreo
Contraloría General de la República, Dirección de Estadística y Censo
Av. Balboa y Av. Federico Boyd
Panamá, Panamá
Tel.: (507) 2104800. (507) 2104870
Fax: (507) 2104801
E-mail: robcastillo28@yahoo.com

BERNARDO GONZALEZ GARCIA
Analista de Muestreo
Contraloría General de la República, Dirección de Estadística y Censo
Av. Balboa y Av. Federico Boyd
Panamá, Panamá
Tel.: (507) 2104800. (507) 2104777, Ext. 4815 o 4816
Fax: (507) 2104801
E-mail: luisqu@contraloria.gob.pa

PERÚ

GILBERTO MONCADA
Jefe
Instituto Nacional de Estadística e Informática
General Garzón No. 654/658, Jesús María
Lima, Perú
Tel.: *(51-1) 4338284 / 4333104*
Fax: *(51-1) 4333159*
E-mail: gmoncada@inei.gob.pe

JUAN CARLOS VALVERDE QUEZADA
Director Ejecutivo
Dirección Ejecutiva de Muestreo y Marcos Muéstrales
Instituto Nacional de Estadística e Informática
General Garzón No. 654/658, Jesús María
Lima, Perú
Tel.: *(51-1) 433 4223 Anexo 185*
Fax: *(51-1) 433 3118*
E-mail: jvalverde@inei.gob.pe

ROSARIO CÉSPEDES KLESCOVICH
Directora Nacional Adjunta
Dirección Técnica de Censos y Encuestas
Instituto Nacional de Estadística e Informática
General Garzón No. 654/658, Jesús María
Lima, Perú
Tel.: *(51-1) 433 4223 Anexo 164*
E-mail: rcespedes@inei.gob.pe

NANCY HIDALGO CALLE
Directora Ejecutiva
Dirección Ejecutiva de Censos Encuestas y Hogares
Instituto Nacional de Estadística e Informática
General Garzón No. 654/658, Jesús María
Lima, Perú
Tel.: *(51-1) 433 4223 (167)*
Fax: *(51-1) 433 6193*
E-mail: nhidalgo@inei.gob.pe

MOISES VENTOCILLA ALOR
Gerente General
Instituto Cuanto
Plaza del Ovalo 203-B
San Isidro
Lima, Perú
Tel. : *(51-1) 4423421*
Fax: *(51-1) 4425460*
E-mail: mventocill@cuanto.org

JESUS GONZA COLMENARES
Gerente de Encuestas
Instituto Cuanto
Plaza del Ovalo 203-B
San Isidro
Lima, Perú
Tel. *: (51-1) 2211675*
Fax: *(51-1) 2211675*
E-mail: jgonza@cuanto.org

HILDEGARDI VENERO
Investigador Auxiliar
Instituto de Estudios Peruanos
Horacio Urteaga 694, Lima 11
Lima, Perú
Tel.: (51-1) 3326194
Fax: (51-1) 3326173
E-mail: hvenero@iep.org.pe

ALBERTO PADILLA TREJO
Consultor
Instituto Nacional de Estadística e Informática
General Garzón No. 654/658, Jesús María
Lima, Perú
Tel.: *(51-1) 850 8142*
E-mail: apadilla_trejo@hotmail.com

DANIEL HOKAMA
Ing. Mecánico
OSINERG
Bernardo Monteagudo 222 Magdalena
Lima, Perú
Tel.: (51-1) 264 0450 anexo 329
Fax: (51-1) 264 0450 anexo 350
E-mail: dhokama@osinerg.org.pe

REPÚBLICA DOMINICANA

LALITO R. VARGAS
Analista II
Banco Central de la República Dominicana
Av. Pedro Henríquez Ureña, Esq. Leopoldo Navarro, Gazcue
Santo Domingo, República Dominicana
Tel.: (809) 2219111 ext. 3109-3110-3111 (oficina) / (809) 6865887 (residencia)
Fax: (809) 6827666
E-mail: lalovargas@hotmail.com

URUGUAY

CARLOS CALVO PELUFFO
Subdirector División de Estadísticas Sociodemográficas
Instituto Nacional de Estadística
Río Negro 1520
Montevideo, Uruguay
Tel.: (598-2) 9008902
Fax: (598-2) 9014653
E-mail: ccalvo@ine.gub.uy

VENEZUELA

ISABEL BAYUELO
Directora de Diseño y Técnicas Estadísticas
Instituto Nacional de Estadística
Av. Bocayá, Edif. Fundación La Salle, Maripérez, piso 4
Caracas, Venezuela
Tel.: (212) 7936083
Fax: (212) 7936083
E-mail: ibayuelo@hotmail.com

RAQUEL FERNÁNDEZ
Estadístico
Instituto Nacional de Estadística
Av. Bocayá, Edif. Fundación La Salle, Maripérez, piso 4
Caracas, Venezuela
Tel.: (212) 7936083
Fax: (212) 7936083
E-mail: raquelf_2000@yahoo.com

CEPAL

JUAN CARLOS FERES
Coordinador MECOVI - CEPAL
División de Estadística y Proyecciones Económicas
Av. Dag Hammarskjold S/N
Vitacura, Santiago, Chile
Tel.: (56-2) 210 2408
Fax: (56-2) 210 2472
E-mail: jferes@eclac.cl

FERNANDO MEDINA
Asesor Regional
División de Estadística y Proyecciones Económicas - CEPAL
Av. Dag Hammarskjold S/N
Vitacura, Santiago, Chile
Tel.: (56-2) 210 2417
Fax: (56-2) 210 2472
E-mail: fmedina@eclac.cl

XAVIER MANCERO
Consultor
División de Estadística y Proyecciones Económicas - CEPAL
Av. Dag Hammarskjold S/N
Vitacura, Santiago, Chile
Tel.: (56-2) 210 2442
Fax: (56-2) 210 2472
E-mail: xmancero@eclac.cl

DAVID GLEJBERMAN
Consultor
Oficina de la CEPAL en Montevideo
Juncal 1305, piso 10
Montevideo, Uruguay
Tel.: (598-2) 9161580
Fax: (598-2) 9161776
E-mail: dglejber@adinet.com.uy

ALFREDO ALIAGA
Consultor CEPAL
11785 Bettsville Drive
Calverton, MD 20705-3119, USA
Tel.: (1-301) 572 0940
Fax: (1-301) 572 0999
E-mail: aliaga@macroint.com

BID

JOSÉ ANTONIO MEJÍA GUERRA
Coordinador MECOVI – BID
1300 New York Avenue, N.W. Washington, D.C. 20577, USA
Tel.: (1-202) 623 3713
Fax: (1-202) 623 3299
E-mail: joseam@iadb.org

Otros Participantes INEI - PERU

Elva Dávila Tanco
Investigador
Dirección Técnica de Demografía
Tel.: (51-1) 433 4223 Ext. 158
E-mail: edavila@inei.gob.pe

Zoraida Castro Angeles
Investigador
Dirección Técnica de Demografía
Tel.: (51-1) 433 4223 Ext. 158
E-mail: zcastro@inei.gob.pe

Bertha Orjeda Túpac Yupanqui
DTD (Demografía)

Clarisa Sánchez Yactayo
Directora Técnica
Oficina Técnica de Informática
Tel.: (51-1) 433 8327
E-mail: clsanchez@inei.gob.pe

Elcira Azañero Terrones
Directora Ejecutiva de Desarrollo de Sistemas
Oficina Técnica de Informática
Tel.: (51-1) 433 4223 anexo 135
E-mail: eazanero@inei.gob.pe

Roberto Mallqui Moreno
Analista de Muestreo
Dirección Nacional de Censos y Encuestas
Dirección Ejecutiva de Muestreo y Marcos Muéstrales
Tel.: (51-1) 433 4223 anexos 185, 340
E-mail: rmallqui@inei.gob.pe

Freddy Castillo Canayo
Analista de Muestreo
Dirección Nacional de Censos y Encuestas
Dirección Ejecutiva de Muestreo y Marcos Muéstrales
Tel.: (51-1) 433 4223 anexo 340
E-mail: fcastillo@inei.gob.pe

Carlos Santur Alberca
Director Ejecutivo de Cartografía y Geografía
Tel.: (51-1) 433 4223 anexo 165
Fax.: (51-1) 433 3118
E-mail: csantur@inei.gob.pe

Arturo Arias Chumpitaz
Director de Metodología
Dirección Ejecutiva de Censos y Encuestas de Hogares
Tel.: (51-1) 433 4223 anexo 168
E-mail: aarias@inei.gob.pe

Eduardo Romero Bolognesi
Asistente Estadístico
CEDRO
E-mail: eromero@cedro.org.pe

Jorge Vilcachagua
Especialista en Calidad de Servicios Eléctricos
OSINERG
Tel.: (51-1) 264 0450 anexo 353
E-mail: jvilcachagua@osinerg.org.pe

Jorge Flores
Especialista en Fiscalización de la Comercialización
de Electricidad
OSINERG
Tel.: (51-1) 264 0450 anexo 339
E-mail: jflores@osinerg.org.pe

Viviana Maldonado García
Asistente Estadístico
CEDRO

Javier Herrera
Consultor IRD

José Llanos
OTI (Informática)